

Appendix A

835 ATTCCTCATGTCGCAATCTCTTA 859

SULF 11

429436

XUS

PRINITION

JCESSION

ERSON

EXORDS

ORICE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

588; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

Db 181 TTTTGACCTTAATTCGACGATGAAGGCTGTGTAAGTCACATGTACACCTGTCG 240

Qy 984 cattactacccctggcaagatcagcaattttctatcactatccatccccaagatgtttt 1043

Db 241 CATTAACCTTGGCAAGATCGACATTTTTCATCTTATCTTCCCAAGATGTTT 300

Qy 1044 tcttgaataaagctagaagagctcctacagcaagagacatttggagagtgagagacc 1103

Db 301 TCTGTATTAAGCTAGAAAAAGTCCCTACAGCAAGAGACATGTGAGAGTGTGAGAAC 360

Qy 1104 atatatatttccaagagagagatgcccacacagatataagaaaaactggagaactgaa 1163

Db 361 ATATATATTTTAAAGAGAGATGCGACAGATTAAGAAAACTGGAGAACTGAA 420

Qy 1164 gagtcagcagatcagcttggcacaagacttgggaatatcgcatttgccttgcctggag 1223

Db 421 GAGTCAGCAGATGATGCTTGGCAAGACTTGGCAATATGCAATGCTTGTGCTGAG 480

Qy 1224 tgcatacattatataatattgttagagagtgctggagtttggagaagatctcagag 1283

Db 481 TCGATTCATTTAATGAAATTTTGTTCAGTCTGAGTGTGAGAAAGATTCATCAGA 540

Qy 1284 agtagaatcagacttgag 1331

Db 541 AGTAGAATCAGTACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 588

1510 bp mRNA linear HTC 03-JAN

4850aa

202 a 99 c 124 g 163 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

588; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

MEDLINE	20530913
PUBMED	11076861
REFERENCE	4 (sites)
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5 (pages 1 to 2710)
AUTHORS	Adachi,J., Aizawa,K., Akabira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bonio,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukushima,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hirooka,T., Hoti,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kozihar,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numataki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,T., Schriml,L., Shibata,K., Shihara,Y., Shinagawa,A., Shiraki,T., Schraml,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 220-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer 15' GAGCAGACAGATCTCCGACTTAATTAAATTCACCCCCCCC 3'. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGCAGACAGATCTCCGACTTAATTAAATTCACCCCCCCC 3'. cDNA was cleaved with BamHI and XhoI. cDNA of size comprised longer than 7 kb was selected before cloning. Vector: a modified plblescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI, 3' end: BamHI. Host: DH109.
FEATURES	location/qualifiers 1..2710 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="MGI:1897080" /db_xref="taxon:10090" /clone="493143IC02" /sex="male" /tissue_type="testis" /clone_id="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult" 1..2536 /note="data source:SPR, source key:Q9P2F2, evidence:ISS homolog to KIAA1395 PROTEIN (FRAGMENT) putative" /codon_start=2 /protein_id="BA30271.1" /db_xref="GI:12855260" /translation="RGRRKHDPICGHGRWSPGSWSLHOPGTFNSSLRLPLSANESSTRLVAVLVMLKNAPELLQRMADALPOLRLDLVLCAAEYGGKAFFERINSITLFKSDMKMARLEALILGTIGARQGVRSRRSPFGQEVNRKRKATROTSDRDYTRKDMEHRAFDGLNLADEASLVLDLETFIVGYMSSEARSISAVLKVVLYSLSSAQSAFLFIGLATOTORAIRAKYLPDLLPEPDLCADCLIRLRCSGISRTIMKHAASVYLVIKRONFEIGHNPRAKMYTMSLSLVGTGNSEHRKSLTYLTVAEEIDILKROSTRAEQVDLMETLNIHTDVTYKMEHQEDPEPMMDMTYLRARYQSSPDL
CDS	

Query Match	16.9%	Score 1156.8	DB 11	Length 2710
Best Local Similarity	68.0%	Pred. No. 2.4e-256		
Matches 1684	Conservative	0	Mismatches 767	Indels 27
			Gaps	4
BASE COUNT	629 a	752 c	771 g	558 t
ORIGIN	RTTUNNAGKAEELGNAEAACAAACVWHAALVAEVLALDESRHLPGVCSEFQWSSN VLEESAIPDDILSPDEFGRCSKNTTELTGLVLEQAGYTFMGILVAPEVCKNLI PLIEAHPRKYKIAAHHKLOAEFTKIMHOSSGWERVETTRVAFYGTGRSDLEDEE YKEPSTIRKLAEISRHLSEFTFERGDVEITIDSNVDSKLDPOAYIQTIVBE HFDVLELDRVYETFERNGRAFLFEFTCTPFDGRHAGLAQHRKTLSTEAFPI KRIIVCHREEVTLVAEVALIEDMOKRKTELAEFTEDOPRPAKILQWLOCSQVPTW OGPLVAQVFLSEIPEDKFLFRNHRKLECKDKCDKCPEDLIRKNKALIGDQKEYHRE ELERYSIRREALQPLLRPLRPLAPLAPSTSLRSMNRSFPRADL"			
3854	aaagttagggcggaatgatataagccgaacccgtttgcatactgcagagcaatgcgaggaatcgg	3913		
1	AAAGGAGCATCGGAACACACATCAACCAATCTGTGGCATGGCCATCTGGGGGTCCCC	60		
3914	tcocctacaataaagccctggcagtttccctccctcaagctcaacgagctggcaagcaaca	3973		
61	TGGCTCCTGAGCTCTGGCACACACATTTCCACAGGACCTTCAAC-AGCAGCTCGTCAGGC	119		
3974	ctaccttttcagcagaatacaagtgcgaacctttgatctgcttactttgggtttcctaaa	4033		
120	TTCTCTCTCTCTGTGAMTCAACCGCGAGCCCTCTGTGTGTCTGTGTGGGTCTGAAGA	179		
4034	atcgacatgaatacaagttctacagaagttttacagatctctcagcttcagcttcgaacaa	4093		
180	ATGCAAGGCAACGCTCCCGACAGCCCTGGGCGTGTGACTGGCCCTGCCCAAGCTGGGGC	239		
4094	ggcattatgactctgcttatactctgctgctgcttgcctttgagatataaagggaataaagt	4153		
240	GACTCTCAGCTACTCTACTCTCTCTGTGGGCTGCTTGAGATCAAGGGGAAGAAAGCCT	299		
4154	ttgaagaaatgaataagcttgcacctttaagaaatcaaaagacatgagcaaaagcttgaag	4213		
300	TTGAGCGGATCAACAGCTCTACGTTCAAGAMGTCTGTGACATGAACGGCGCTGGAGG	359		
4214	aagctattcttggagcatalagctgcacagcaagaatgttcacgaggaacgagagacagc	4273		
360	AAGCATCTCTGGGCAACATTTGGGGCGGGGACAGAGATGTTGGGGGAGTGGGGAGA---	416		
4274	tcgagagaagcccatctggaagtgaccttttgaagtccaagaaatttgagtgaggaag	4333		
417	-----GAAGTCCATTTTGGAAACCAAGAAATGTTGCTGGGGGAGA	458		
4334	atactactactgcgtcaaaaacagagaagcttgcacaaatcaagacagaattgaac	4393		
459	GTGCCACACACTGAGACAGACACTCAGACCGGTGTGACAAACCAAGATTAATGGAAC	518		
4394	acgaagcaactgattgattggaacacttgctctacagaagaacaaactaatcaattagatacat	4453		
519	ATGAAGCTCTGTGTGACGGAACACCTGGGACAGAGAGCGCCAGCTGTGTTCTGACACAGC	578		
4454	tagaagattgtttcaagccgtttctctgtaaggaatccaagaagagaattcttggtagag	4513		
579	TGGAAACCATGTGTGACAGACAGTATGCTGTCCGAGGCCCCGTGAAGGCACTTCTGATGGCG	638		
4514	tgctaaagtgctactaacacagcattgctcttlaaccaaagtgcagtttatctataaacact	4573		
639	TGCTGAAGATTGTCTTACAGTCTTTGGAGGGCCACAGTAGCCCTGTCTCTGAGCATG	698		
4574	gttttgctacaagaagaccttggtttcaaaagtttctgaaactcttatttgaagaaga	4633		
699	GGCTGGCAACCCAGGGGCGCTGTCTTCAAGATTTCGGACCTCTTTTCGAGGAACACA	758		
4634	cagagcagtgctgtaattatgaactcgaagcttctccgaacacttgtagaagtagacgtgta	4693		
759	CGGACCTTTGTGCGACCTGTGCTGAGACTTCTTGCGACACTGTGGGACCGGCAATCGACA	818		
4694	caatacggctcaacccacatgctctcccttactactaatatgagcaaaactttgaattg	4753		

Db 819 CCATCCGATGACGACGAGGCCCTCCCTACCTGCTTATGCCGCAACCTTGAGATTG 878
 4754 ggaataacttggcagggttaaaatgcaggttacccaatgtaactaactccttggtgga 4813
 879 GCCATTAACCTTGGCCCGTGTGATGCTGGTACACATGCTCTCTGCTCCCTTGTGGGA 938
 4814 catctcagaatttaagaagaattcttaagcgttctctaagaactataatgacataig 4873
 939 CAACTCAGAACTTATGATGAGAGCATTGAGAAAGTCCCTCAAGACCATCTGACCTACG 998
 4874 ctgaagaagatctggaattgaggaacaacattcctgatacaggtccagatctggtt 4933
 999 CAGAGGAGGACATGAGGCTGAGGACAGCACTTCGCTGAGCAGTCCAGACCTCATGT 1058
 4934 tcaatctccataatgattcttctgatactgtgaatgaagaaggaacacagagatctg 4993
 1059 TCAACCTGACATATCTCTGACAGACAGGTGAAGATGAAGAACACAGAGAGACCTTG 1118
 4994 aaatgtgatgatctaatagtacagaattgccaagggttaaccagactctccaga--gc 5050
 1119 AGATGCTCATGAGCTCTCATGACAGAAATGCTGGGGATACCAAGGCTCTCCACACTGC 1178
 5051 gattgacctggttgcaagaatgcaaggcaagcaactcaagaacgaaacatcatgtgaag 5110
 1179 GCGTGACATGCGCTGACAGAACTGCTGCAAAACATGCAAGAGCTGGCAATCATCATGAG 1238
 5111 ctgcaagctgtctgatacactcagcagcactgtgtctgaattatgtgaatgtgtgaag 5170
 1239 CTGCCAATGATGATGCTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAG 1298
 5171 accggaatatactctctggtgagatgtgaacattcagaattatcctaattcttag 5230
 1299 ACAGTGGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1358
 5231 aagaactcgcggtctcagaatgtgtgtaactccaagatgaagaagatctgtctctgaa 5290
 1359 AGGACTCTGCCATCTCTGATACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAG 1418
 5291 aatacttactgagtcagcagacttctggatctactggaacagcagcgtctctctctta 5350
 1419 AGAATTTACAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1478
 5351 tggctgcatatgaagaacagtttaagaagtttaacaagttactatctcttaattgaag 5410
 1479 TGGGTGCTGCTGATGAACGCGTGAACGATCTACAAAACCTTATCCCATCTCTGAG 1538
 5411 ctaatcggatgcaagaactatccacaatctatgtaactgaactcaagaagcattcaga 5470
 1539 CCCACAGAGACTACAAAGAGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTG 1598
 5471 aatgtgtcatcagactgactgctgagcgagtggttggcagcactatttgggtt 5530
 1599 AATTAATGACACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1658
 5531 ttatagaaccaagttcggggaatttgaatgaagaagaatttggttacaagaagcctgaa 5590
 1659 TCTATGGCACAGATTGCTGATGCTGATGAACAAGACTTGTGATGAAGAAGACCGTCAA 1718
 5591 taacaaactgcaagatatctcagatctgaggaatttcaaggaaagaatttggag 5650
 1719 TCCACAGAGCTTCCAGAGATCTCCACCGCTGGAGAGAGTCTTATACGGAAGGTTGGGG 1778
 5651 agagatgtgtgaataatcaagaagcctatctcgttagacaagtgtaataatgaactcta 5710
 1779 ATGATGTGTTAGATCATCAAGATTCTAACCCAGTGGACAGATCCAAAGCTGGACCCAC 1838
 5711 acaaggcatataatcaattactatgtgagccaacttgcacacatatagatgaag 5770
 1839 ACAAGGCGTACATACAGATTAACCTATGTGAGACACATTTGACACATTATGAGCTCAAG 1898
 5771 acaagatcaactatttgcagaaaaataataatctctcgatcgatctatgtatgaacct 5830

Db 1899 ATCGGTGACACTTCTTCATGCGAACTATGAGCTGCGGGGCTCTCTCTTCTGACACCT 1958
 5831 ttaacttaagtgcgctgcccattgggaactlcatgaactlcaaaagaagaccattc 5890
 1959 TCACACAGATGAGAGCTGGAGAGCTGGAGAGTGGCCGAGACAGCAAGAGACGAGCTGC 2018
 5891 tgaactcgtlcatatgcttctcttataataaacaagggtcaatgtlcatcactaaagaag 5950
 2019 TGAGCAGGAGAGATGCTTCTCTCTCATCAAGACAGCATCGATGCTGCTGCTGCTGCTG 2078
 5951 agatcatcttaacccaacttgaatgttctatgtatgaagacatgcagaaagaacagagat 6010
 2079 AGACATGCTGACACCAAGTGAAGTGGCTATGAGACATGCAAGAGAACCCGGAGAC 2138
 6011 tgcatttgcacacatacagatcccgagaccaccaaaatgtcttaagatgtactcagg 6070
 2139 TGCCCTTGGCAGCCAGCAGAGACCTCCAGATGCCAAGATGCTGCAAGTGTCTCCAGG 2198
 6071 gactgtgaagcaccagatgaatcaagggtcttgaagatgtgccaggttttctgtctg 6130
 2199 GTTCTGTGGACCACTGCAACCAAGGCTCCCTTGCAAGTGGCCAGAGTGTCTTGTGAG 2258
 6131 aaatcactagtagcccaagctcttcaagatcatataataactlgaactctgtctaaag 6190
 2259 AGATCCAGAGATCCCAAGCTCTTCGACATCACAACACTCGGCTCTCTTCAAG 2318
 6191 atttactaaagggttgaagatgcttaagaataaagaagcttaattggccggttc 6250
 2319 ATTCTGCAAAAAGTGCAGAGATGCACGTAGAAAGACAGGCCCTGATTGGCCCAAGAC 2378
 6251 aaaaagatatacaagggaattg--gggaactatcttcgctttaaaggccctacagc 6308
 2379 AGAAGAGATACCCAGGAGCTGAGCTGACATATGAGCCGCTGGGAGGCTCTGACAC 2438
 6309 cctagatcaagaagatc 6326
 2439 CTCTGCTTACCAACGTC 2456

RESULT 2
 LOCUS BC018599 1311 bp mRNA linear HTC 03-DEC-2001
 DEFINITION Mus musculus, clone IMAGE:3596745, mRNA.
 ACCESSION BC018599
 VERSION BC018599.1 GI:17235789
 KEYWORDS HTC.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (Bases 1 to 1311)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK
 COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: villalondm.tmc.edu
 Villalon, D.K., Luna, R.A., Hale, S.M., Huylk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.
 clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRK Plate: 16 Row: e Column: 2
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis
 This clone has the following problem: no 5' EST match.

FEATURES

source	1. .1311
--------	----------

```
"organism":"Mus musculus"  
"db_xref":"taxon:10090"  
"clone":"IMAGE:3596745"  
"tissue_type":"Mammary tumor, C3(1)-tag model,  
ductal carcinoma, 5 month old virgin mouse."  
"clone_1ib":"NCI CGAP_Mam6"  
"lab_host":"DH10B"  
"note":"Vector: pCMV-SPORT6"
```

```

/Note="Vector: pCMV-SPORT6"
BASE COUNT      363 a      309 c      333 g      306 t
ORIGIN

```

Query Match	15.3%;	Score 1044.4;	DB 11;	Length 1311;
Best Local Similarity	87.9%;	Pred. No. 1.9e-230;		
Matches 1151;	Conservative	0;	Mismatches 156;	Indels 3; Gaps 1;

OY	3832	tgatagccacatgatatgatgaagaagagagagagatgttgtaagccagacctgtgc	389
Db	1	tgtatagcccttgatgatattgacagtgtagtgtagtgagcattgataaacccagacggtggcc	60
OY	3892	atgacaatcgcagggagacatcgtgctccctcaacaaacaagagctgagcattccctccaaag	395
Db	61	atggcaattggcaggaacatccgctccctcagctgacaaagacctggcattttcttcacag	120
OY	3952	tcaagagatggcagggcaacacacactaccttttcagcagaatcaagtcgaagcctttgatc	401
Db	121	tcmaagatggcccgccacacacacaccccttctcacacagtcgaagcagagaccccttgatc	180
OY	4012	tgtcacttggattctccaanaatcgagatgaaacacatttcaagaagtgtttaagat	407
Db	181	tgtctgcttttgggtttctcamaaattgacagatgaataacggcttctgcagaaattggttaacagat	240
OY	4072	ctctcagctctgcagctaaacccgctattagatcgtcttattctctgtgtctgctt	413
Db	241	ctctcagctctgcagctaaacccgctattagatcgtcttattctctgtgtgtatcttgccttt	300
OY	4132	gagatataagggnaaanaagtgcttgaacgaatgaaatagccttgaaccttaagaaatcaaa	419
Db	301	gagtcnaaggaanaaaaggtattttgaaacgaatgaatagtttaacttttaagaatacaaa	360
OY	4192	gacatgagagagaagccttgaagaagcattctcttggagagctaatgtccagcgcaagaatg	425
Db	361	gacatgagagagaagccttgaagaagcattctcttggagagcattgtccgctccagcgcaagaatg	420
OY	4252	gtacggcgaaagccgagagacagctcgagagagaagcccaatcgaaagtgcctttgaatcaa	431
Db	421	gtacggcgagacccagggcacctcgaaagagagcccatcttgaaagccgctttgggacccaa	480
OY	4312	gaaaatttgaggtgagagaaagatatgactcactggtgctcaaaaacagagaagccttgac	437
Db	481	gaaaacctgcggtgagaaagagacatgactcactggtgctcaaaaacagagaagccttgac	540
OY	4372	aaatcaagagagagatggaacacgaagacatgatatgagaaacctggtctacagaaga	443
Db	541	aaatcaagagagagatggaacacgaagacatgatatgagaaacctggtctacagaaga	600
OY	4432	aaactaatcatattagataactaatgagaattggtgttcagagccgttctgtaacgaatcc	449
Db	601	aaactaatcatattagataactaatgagaattggtgttcagagccgttctgtaacgaatcc	660
OY	4492	aaagagagcattctgtgtgagtgctaaagtgctactaaacagaaatggcctttaaaca	455
Db	661	aaagagagcattctgtgtgagtgctaaagtgctactaaacagaaatggcctttaaaca	720
OY	4552	agtgaggttatctacaacaacgttttgtctacacagagagccttgglttccaagtctct	461

Db	721	ACMGAGTATCTGTGACGACGCTTGGCCACGACGAGAGCCCTGGTCCAAAGTTTCT	780
QY	4612	gaactcttatattgaaagaagaacagacgagtgctgtatattatgcctcagcgtcttcoga	4671
Db	781	GAGCCTCTGTGTTGAGAGAAAGACAGAGCGAGTGTGGGATCTGTGGCTCCGGCTTTCCTT	840
QY	4672	caacttaagcagtagatctgctgtaacaatagtgtaaccccaagtcgtcccttaactacta	4731
Db	841	CATTGCAGTAGACGACATCGTACATTCGGTGTGCACGCTAGTGCCCTCTTACCTCCTC	900
QY	4732	atgagagcaaaactttgagatttggaaataaacttgcagaggtttaaaatgcaagtlaccatg	4791
Db	901	ATGAGGCGAACAATTGAGATCGGAGATTACTTTGGCAGAGTGAAATGCAAGTAGACGATG	960
QY	4792	tcaactatctctctgtgtgggcacactctcaaatltaatgaagaattctctaaagcgtct	4851
Db	961	TCACGTGCTCTTTGGTGGGTACGTCTCAGAAATTTTAAAGAAGATTCTTGAAGACGCTCT	1020
QY	4852	ctaaagacatactgacatatgctgaagaagaatctggaaattgaggaaacaacttctct	4911
Db	1021	CTAAAACATATTGTACATATGCTCAGAGAGAGATTAGAAATTAGGGAACCACTTTTCTT	1080
QY	4912	gacacgttcacgagatctggttttcaactcatalgatctcttcgtatactgtgaanaag	4971
Db	1081	GATTCAGGTCCTCAAGATTGTGGTCTTCATATCCATATGATGATCTTTGTACACTGTTAAGATG	1140
QY	4972	aaggaaccccgaggaaagatctcgbaaatgttgtatgtgtctcaatgtlaaagaattgccaaggt	5031
Db	1141	AAGGAGCATTCAGGAAGACCCGAAATATGCTGATTAATCATATGTACAGAAATCCCAAGGGC	1200
QY	5032	taccagacctctccaga---gcgatttgacctgtgttcgaagaatgagcgaaagacactca	5088
Db	1201	TACCAAGAGTCTCCAGACACTGGCCTGAACTGTGTCAGAAATATGGCTGGCAAGACACTCT	1260
QY	5089	gaacgaagaacatcatgctggaagctgcacagatgctctagctccactcagcagc 5138	
Db	1261	GAAACCAACATTCATGCGGAAGCTGCTCATAGTCCCTGGTGCATCTCCGCGAGC 1310	

RESULT	3
LOCUS	BF309712
DEFINITION	BF309712 898 bp mRNA linear EST 21-NOV-2000 601891854F1 NIH_MGC_17 Homo sapiens CDNA clone IMAGE:4137732 5', mRNA sequence.
ACCESSION	BF309712
VERSION	BF309712.1 GI:11257158
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 898) NIH-MGC http://mgc.ncl.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) unpublished (1999)
AUTHORS	Contact: Robert Strausberg, Ph.D.
TITLE	Email: cgapbs-remail.nih.gov
JOURNAL	Tissue Procurement: ATCC
COMMENT	CDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at: image.llnl.gov Plate: LHCMI045 row: k column: 13 High quality sequence stop: 132.

```
FEATURES
source
1. 898
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:413773"
/tissue_lib="NH_MCC_17"
/tissue_type="Thadomyosarcoma"
```

/lab_host="DH10B (phage-resistant)"
 /note="Organ: muscle; Vector: pGB7; Site: 1: EcoRI;
 Site: 2: XhoI; cDNA made by oligo-dT priming.
 directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(C). Size-selected
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 276 a 182 c 228 g 212 t
 ORIGIN

Query Match 9.9%; Score 678.8; DB 10; Length 898;
 Best Local Similarity 92.5%; Pred. No. 6e-146;
 Matches 757; Conservative 0; Mismatches 57; Indels 4; Gaps 4;

OY 4152 gttgaacgaatgaatgactgacatttaagaatcaaaagacatgagagaagcttga 4211
 DB 2 GTTTGAACGAATGAATGACTTGAACCTTAAGAAATCAAAAGACATGAGACCAAGCTTGA 61
 OY 4212 agaagcattcttggagccttagtgccaggaagaatgctcggaggaagccgagaca 4271
 DB 62 AGAAGCTATTCTTGGAGCATAGGTGCCAGGCAAGAAATGTACGGCGAAGCCGAGGACA 121
 OY 4272 gctgagagaagcccatctggaagtgcttggagagcaagaatlttgagtgagagaa 4331
 DB 122 GCTGAGAGAGAGCCCATCTGGAATGCTTTGGAAGTCAGAAATTTGAGTGGAGAGAA 181
 OY 4332 agatatgactcactgagcgtcaaaaacacagagaagcttgacaatcaagacagaagattga 4391
 DB 182 AGATATGCTCCTACCTGGGCTCAAAACACAGAGAACCTTGACAAATCAACAGACAGATTGA 241
 OY 4392 aacagaagcactgattatgtaagaacctggtcacagaagaacactaactaatttaagttac 4451
 DB 242 ACACGAGACACTGATTGATGGAACCTGGCTACAGAGACAACTTATCTATTTAGTAC 301
 OY 4452 attagagattgttcaagacgttctgtaacggaatccaaagagacattctgtgtg 4511
 DB 302 ATTAGAGATTG-TGTTACAGACCGTTTCTGTAAAGCAATCCAAAGAGAGCAATTCTGTGG 360
 OY 4512 agtctaaagtgtctactacacagatgctgtcaacaaagtgcattatctacacaa 4571
 DB 361 AGTCTAAAGTGTCTACTACACAGCATGCGCTGTAAACAAAGTGCAGTTATCTACAAACA 420
 OY 4572 ctgttttctcacagagaagcctgtgttcaaaagtttctgaactcttatttgaagaaga 4631
 DB 421 CTGTTTCTACACAGAGCGCTTGTTTCAAGTTTCTGAACTCTTATTGAGAGAGA 480
 OY 4632 gacagaagcagtgctgattatgctcgaagctctcgcgaacgttagcagtagacatcg 4691
 DB 481 GACAGAGCAGTGTGCTGATTATGCTCAGGCTTCTCGACACAGTGAAGTACATCGG 540
 OY 4692 tacaatcagtcacaccccaagtcctcccttactactaactaataagggcaaaacttggat 4751
 DB 541 TACATACAGGATTCACACGCGACGTGCTCCCTTACCTACTATATGAGGCAAAAC-TTGGAT 599
 OY 4752 tgggaataacttgcacaggtttaaataatgacgaatgacatgacatcccttcgtggtg 4811
 DB 600 CGGGAATTAATCTTGGCAGGGTTAAATGACGTTACATATGCACTATTCCTCC-TGGTGG 658
 OY 4812 cacatctcagaattttaaagaattcctlaagaagcttctcctlaagaactatattgaac-t 4870
 DB 659 CACATCTCAGGAATTTTAATGAAGATCTTAAGAGCTTCTTAAGACTATATTATGACACT 718
 OY 4871 atgtcgaagaagatctggaatltggaggaacaaacttcccgatcaggtccagatctgg 4930
 DB 719 ATGCTGGAGGAATCTGGCAATGGGGGAGACCTTCGCGGATCAAGGCCAGAGATCGG 778
 OY 4931 ttccaatccataatgattcttctgatactgtgaac 4968
 DB 779 GTTCACACCACTATGAGTTTCTGTGCTACTGTGAGAGA 816

RESULT 4
 BC018075 1873 bp mRNA Linear HTC 06-DEC-2001
 LOCUS Homo sapiens, clone IMAGE:4792506, mRNA.
 DEFINITION BC018075
 ACCESSION BC018075.1 GI:17390162
 VERSION
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1873)
 Strausberg, R.
 Direct Submission
 Submitted (03-DEC-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-rt@mail.nih.gov
 Tissue Procurement: Miklos Palcovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
 Yoshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILND)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mdc@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILND at: <http://image.llnl.gov>
 Series: IRK Plate: 32 Row: b Column: 21
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Similarity but not
 identity to protein
 This clone has the following problem: frame shifted.

FEATURES
 source location/Qualifiers

1..1873
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4792506"
 /tissue_type="Brain, hypothalamus"
 /clone_id="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"
 BASE COUNT 604 a 317 c 365 g 587 t
 ORIGIN

Query Match 9.9%; Score 678; DB 11; Length 1873;
 Best Local Similarity 100.0%; Pred. No. 1.2e-145;
 Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5498 agggagtggttggcaccactatttcgtgtggttttaaggaaccaagttcggggaattgg 5557
 DB 1178 AGCGGATGTTTGGACACTATTTCGTGTTGGTTTATATGAAACCAAGTTCCGGGATTTGG 1237
 OY 5558 atgaacaagaatttgttcaagaagcctgcaataaaccacttgcagagatatctcaca 5617
 DB 1238 ATGAACAAGAAATTTGTTTCAAGAGACCTGCAATACCAAACTGCGAGATATCTACACA 1297
 OY 5618 gatgtgaggaatttccggaagaagatttgggagagatgtgttgaagtaactcaagaact 5677
 DB 1298 GATTGGAGGATTTTACGGAGAAAGATTGGAGAGGATGTGTTGAATCAATCAAGAGACT 1357
 OY 5678 ctatcctgtagaagaagtgttaaatagatccttaacaagaagcatatattcagattactatg 5737
 DB 1358 CTATCTCTGAGCAAGTGTAAATTTAGATCCCTATACAGGCAATATATTCGATTACCTATG 1417

```

QY 5738 tggagccactatgacacatatgatgaagacagacatcaactatctgcacaaaatt 5797
      |||
Db 1418 TGGAGCCACTACTTGGACACATATGATGAGAGACACAAATCCTATTGGACAAAATT 1477
QY 5798 acaactctgcgtacatcaatgactgacacaccttactttagatgagcgctggccatggg 5857
      |||
Db 1478 ACAATCTCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1537
QY 5858 aactctgacacacacacacacacacacacacacacacacacacacacacacacacacac 5917
      |||
Db 1538 AACTCTGATGACAAATCAAAAGAGACCATCTGACTGACTGACTGACTGACTGACTGACT 1597
QY 5918 ttaaaaaagggtcaatgacacacacacacacacacacacacacacacacacacacacac 5977
      |||
Db 1598 TTAATAACAGGCGCAATGTGCTCCTCATTAAGAGAGATCATCTTAACACCAATGAAATGG 1657
QY 5978 ctatggagacatgcagaaaaagacagagatgtgacattgcacacacacacacacacac 6037
      |||
Db 1658 CTATTGAGGACATGCAGAAAAAGACAGAGATGTGGCATTTGCAACACATCAGGATCCCG 1717
QY 6038 cagagcccaaatgctcagatgctcagagatctgtaggcacacacacacacacacacacac 6097
      |||
Db 1718 CAGACCCCAAAATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1777
QY 6098 ggccttggagagttgcacaggtttctctgctgaataactagtgacccaagctctca 6157
      |||
Db 1778 GGCTTTGGAAAGTTGCCAGGTTTCTGTGCTGAATACCTAGTGAACCCAAACCTCTTCA 1837
QY 6158 gacatcataataaactgc 6175
      |||
Db 1838 GACATCATATATTAACATGC 1855

```

```

RESULT 5
LOCUS BC018076 1873 bp mRNA linear HTC 06-DEC-2001
DEFINITION Homo sapiens, clone IMAGE:4792507, mRNA.
ACCESSION BC018076
VERSION BC018076.1 GI:17390164
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1873)
Strausberg, R.
Direct Submission
Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web Site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcdexax1.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 32 Row: B Column: 20
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
 This clone has the following problem: frame shifted.
 Location/Qualifiers

```

source 1..1873
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4792507"
/tissue_type="Brain, hypothalamus"
/clone_id="NIH_MGC_96"
/lab_host="DH10B"
/note="Vector: pBluescript"
BASE COUNT 604 a 317 c 365 g 587 t
ORIGIN
Query Match 9.9%; Score 678; DB 11; Length 1873;
Best Local Similarity 100.0%; Pred. No. 1.2e-145;
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5498 agcggatgttggcaccatcttcgtgtgtgtttaaagaacagtcgggattcg 5557
      |||
Db 1178 AGCGGATGTTGGACACCTATTTCGTTGTTTATGAAACCAAGTTCCGGGATTTGG 1237
QY 5558 atgaacaagaattgtttacaaggagcctgcacataccaactgcagagatctca 5617
      |||
Db 1238 ATGAACAAGAATTTGTTTACAAGGAGCCTGCAATTAACCAACTTGCAGAGATATCTACA 1297
QY 5618 gattggagagattttacggagaaagatttggagagatggttgaagtaataaagact 5677
      |||
Db 1298 GATTGGAGGATTTTACGGAGAAAGATTTGGAGAGATGTGGTAATCAATCAAAAGACT 1357
QY 5678 ctaatccttgaacaagtgtaataatagatccctaacaagacatatactacatactg 5737
      |||
Db 1358 CTAACTCCTAGACAAAGTGAATTAATTAATGATCCTAACAAAGCATATATTCAGTTACTATG 1417
QY 5738 tggagccactcttgacacacatgagatgagagagacagacatacctattgcacaaaatt 5797
      |||
Db 1418 TGGAGCCACTACTTGGACACATATGATGATGATGATGATGATGATGATGATGATGATGATG 1477
QY 5798 acaactctgcgtacatcaatgactgacacaccttactttagatgagcgctggccatggg 5857
      |||
Db 1478 ACAATCTCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1537
QY 5858 aactctgacacacacacacacacacacacacacacacacacacacacacacacacacac 5917
      |||
Db 1538 AACTCTGATGACAAATCAAAAGAGACCATCTGACTGACTGACTGACTGACTGACTGACT 1597
QY 5918 ttaaaaaagggtcaatgacacacacacacacacacacacacacacacacacacacacac 5977
      |||
Db 1598 TTAATAACAGGCGCAATGTGCTCCTCATTAAGAGAGATCATCTTAACACCAATGAAATGG 1657
QY 5978 ctatggagacatgcagaaaaagacagagatgtgacattgcacacacacacacacacac 6037
      |||
Db 1658 CTATTGAGGACATGCAGAAAAAGACAGAGATGTGGCATTTGCAACACATCAGGATCCCG 1717
QY 6038 cagagcccaaatgctcagatgctcagagatctgtaggcacacacacacacacacacacac 6097
      |||
Db 1718 CAGACCCCAAAATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1777
QY 6098 ggccttggagagttgcacaggtttctctgctgaataactagtgacccaagctctca 6157
      |||
Db 1778 GGCTTTGGAAAGTTGCCAGGTTTCTGTGCTGAATACCTAGTGAACCCAAACCTCTTCA 1837
QY 6158 gacatcataataaactgc 6175
      |||
Db 1838 GACATCATATTAACATGC 1855

```

RESULT 6
 AK004816 2694 bp mRNA linear HTC 19-JAN-2002
 LOCUS AK004816
 DEFINITION Mus musculus adult male lung cDNA, RIKEN full-length enriched
 library, clone:1200017A24:homolog to BA165F24.1.1 (MOVIE PROTEIN
 SIMILAR TO DROSOPHILA CG6630 AND CG11376, KIA1058, RAT TRG
 (ISOFORM 1)) (FRAGMENT), full insert sequence.
 ACCESSION AK004816


```

QY 5287 ggaataacttactgaagcaggaacttgggtggaacttgcgaacaagcagctgtctctc 5346
    || || || || || || || || || || || || || || || || || || || || ||
Db 421 GGTGGTACTTACTGAGAGTGCTGGTGGCTCTGCGACAGGCTGGAGCTTTC 480

QY 5347 tctatgctgcatlbtatgaagcaatlaatgaagttacaagaagttacttactcat 5406
    || || || || || || || || || || || || || || || || || || || || ||
Db 481 AACAGCGAGGCTTGTAGACAGCGGTATATGAAGTCTCAAGCTGTCTATCCCTACTG 540

QY 5407 gaagctaacgagatgacaaagaactatccacaattcgtttaaacttcaagaagcaltc 5466
    || || || || || || || || || || || || || || || || || || || || ||
Db 541 GAGGACACAGAGATTTCCGAGAGTGCCTCCTCCTCAGCAACAGCTCAGAGGCTTTC 600

QY 5467 agcaaatgttccatccagatctgctgggagcgagatgttggaacatttttctgtt 5526
    || || || || || || || || || || || || || || || || || || || || ||
Db 601 GATTAACATCATCAACAG-----GACCTTAAGAGGATGTTTGGAGCTTCTCGAGCT 654

QY 5527 ggttttatggaacaagttcgggatttgaatgaacaagaattgtttacaagagacct 5586
    || || || || || || || || || || || || || || || || || || || || ||
Db 655 GGTTCCTACGATGCCGATTTGGGATTTGATGAGAGGAGTGTGTACAGGAACCC 714

QY 5587 gcaataaccaaacctgcagagatatctcacagattggagagattttacgagaaagattc 5646
    || || || || || || || || || || || || || || || || || || || || ||
Db 715 GCAATTCAGAAAGCTCCGAGATCTCACATAGACTAGAGGATTTATGGCCAGTGTTC 774

QY 5647 ggaagagatgtgtgaagtaatacaagacttaactcctgtagacaagtgtaaatagat 5706
    || || || || || || || || || || || || || || || || || || || || ||
Db 775 GGTGACAGATTTGTGGAAGTGTATAAAGACTTACTCCAGTGCACAAACCAAGTTGGAT 834

QY 5707 cctaaacaaggcalatatacgaattacatltgagacatacttgcacatatgagatg 5766
    || || || || || || || || || || || || || || || || || || || || ||
Db 835 CCTAACAGGCTTACTACATCACTTTTGGAGCCTTACTTGTATGATGATGAGATG 894

QY 5767 aagagcaaatcactatttgcacaaaataataactctgtctgaatttactgtaca 5826
    || || || || || || || || || || || || || || || || || || || || ||
Db 895 AAGAGCGGGTGACTTACTTGCAGAAAGATTCAACCTCCGAGGTTCTATGACACCC 954

QY 5827 cccctactttagatgagcgtgccatgagggaacttcatgaacaattcaagaagaagacc 5886
    || || || || || || || || || || || || || || || || || || || || ||
Db 955 CCTTCACCCCTGAGAGGGAGACCCCGGGCGAGCTTCAATGAGCAACCCGAGAAACCC 1014

QY 5887 attctgaactgctcactgacttcccttatattaaacaagggtcaatgtactcataa 5946
    || || || || || || || || || || || || || || || || || || || || ||
Db 1015 GTGCTCACACACCATCGACGCTTCCCTCACTCAAGACACAGAGATCCGACGACGAAA 1074

QY 5947 gaaagatcatccttaacacaattgaattgtctattgagcactgcagaaataaacacag 6006
    || || || || || || || || || || || || || || || || || || || || ||
Db 1075 GAGGAGTTCTGTTTACTTCGATTCGATGAGTTGCCATTGAAGATATGAGAAAGAACCCCTG 1134

QY 6007 gaattgcatltgcaacacatcagatcccgacagccccaaatgcttcaagatgtactc 6066
    || || || || || || || || || || || || || || || || || || || || ||
Db 1135 CAGTTAGCCGTGCGACACACACAGAGGCCCTCGATGCAAGAGTGTCTCAATGTACTG 1194

QY 6067 caggagatctgttaggcaacacagtgatcaggggcccttggaaagttgccagtttctgt 6126
    || || || || || || || || || || || || || || || || || || || || ||
Db 1195 CAGGGCTCTGTAGGAGCAGTGTAAATCAGGACCACTGTGGGTGGCCCAAGTGTCTTG 1254

QY 6127 tctgaataccttagtaccacaagctcttcagacatcctaataactggagctgtgtt 6186
    || || || || || || || || || || || || || || || || || || || || ||
Db 1255 GCTGAATTTCCAGCTGACCCCAAGCTCTACGACATCAACAACAGCTGAGTGTGTCTC 1314

QY 6187 aaagaatttactaaaggltgtaagatgctttaaagaaaaa laagagcttlaattgggacg 6246
    || || || || || || || || || || || || || || || || || || || || ||
Db 1315 AAGAGTTTCATATGTCGATGCGGAGAGCCCTGGAAGAAACAGGCGATCATCACCGCA 1374

QY 6247 gtccaagaagatcatcaagggaattggga--aaactatctgccttaagaagacctta 6304
    || || || || || || || || || || || || || || || || || || || || ||
Db 1375 GAGCAGCGGGAGTACACACAGAGACTGAAAGAAACTACAAACAACTGAGAGAGCTTC 1434

QY 6305 cagccctagat--cacagaagaagtcctcactatccaaagcaagatattgctgtccctgc 6362
    || || || || || || || || || || || || || || || || || || || || ||
Db 1435 AGGCCCATGATTTGAGCGGAAAAATCCAGAGCTTACACAGGCCCATATTTACAGATTGACAGT 1494

```

```

QY 6363 caca 6366
    || ||
Db 1495 CAGA 1498

RESULT 7
LOCUS BI088758 795 bp mRNA linear EST 20-JUN-2001
DEFINITION 602851347F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:493330 5',
mRNA sequence.
ACCESSION BI088758
VERSION BI088758
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 795)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1013 row: m column: 11
High quality sequence stop: 640.

FEATURES
    source
        1..795
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:493330"
            /clone_1lb="NIH_MGC_10"
            /cell_line="MGC36"
            /lab_host="DH10B"
            /note="Organ: cervix; Vector: pCMV-Sport6; Site_1: NotI;
            Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
            Average insert size 1.5 Kb. Library prepared by Life
            Technologies."
BASP COUNT 256 a 155 c 189 g 195 t
ORIGIN
Query Match 9.5%; Score 651.4; DB 10; Length 795;
Best Local Similarity 96.0%; Pred. No. 1.3e-139;
Matches 765; Conservative 0; Mismatches 21; Indels 11; Gaps 9;

QY 4074 ctcaagcttgcagctaaacgcg-c-tatagatcgtcttactctgtgtgtctgttg 4132
    || || || || || || || || || || || || || || || || || || || || ||
Db 1 CTCAGCTTTCACACTAACCAGCGGATATGATCTGCTTATCTGTGTCTGCTTGG 60

QY 4133 agfataaaggaaaaaagtggttgaacgaatgaatgcttgcacctttaaagaataaag 4192
    || || || || || || || || || || || || || || || || || || || || ||
Db 61 AGTATTAAGGAAAAAGTGTGTAAGCATATGACTTGAACCTTTAAGAAATCAAAAAG 120

QY 4193 acatgagagcaagcttgaagaagctatctcttggagacatagtgccaagcaagaatgg 4252
    || || || || || || || || || || || || || || || || || || || || ||
Db 121 ACATGAGAGCAAGAGCTTTGAAGAAGCATATCTTGGAGCATAGTGTCCAGGCAAGAAATGG 180

QY 4253 tacgcgaagccgagagcagctcgagagaagcccatctggaagtgcttggaaatcaag 4312
    || || || || || || || || || || || || || || || || || || || || ||
Db 181 TACGGCGAAGCCGAGGACAGCTCGAAGAACCCCATCTGGAAGTGCTTTGGAAGTCAAG 240

QY 4313 aaaattggagggtggaagaagatatgactcactggtgttaaaaacagaagaagcttgca 4372
    || || || || || || || || || || || || || || || || || || || || ||
Db 241 AAAATTGAGGTGAGGAAAGATATGACTGTGGGTCAAAAACACAGAGAGCTTGACA 300

QY 4373 aatcaagagcagagattgacacgaagcactgtattgatgaaacctgtacagaagcaa 4432

```


polyA_signal 1031..1036
 /note="putative"
 polyA_site 1051
 /note="putative"
 BASE COUNT 337 a 207 c 199 g 308 t
 ORIGIN

Query Match 9.5%; Score 646; DB 11; Length 1051;
 Best Local Similarity 82.6%; Pred. No. 2.4e-138;
 Matches 876; Conservative 0; Mismatches 160; Indels 24; Gaps 11;

```

Qy 5763 gatgaaggacgaatcaccttcttcgacaaaataacatctcgtcgtatcgtactg 5822
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 GATGAAGACAGAAATCACTTATTTGACAAAAATTATATCTCGCGCTTCACTACTG 61
Qy 5823 tacaccccttacttaagtgcgcgtgccatgagggaacttcatagaacaattcaaggaa 5882
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 CACACCCCTTCACTTAAATGAGCCGCTGCATGAGGAGCTTCATGAAATTCAAACGGAA 121
Qy 5883 gaccatctgactacgtctcatgctcttctatattaaacaagggtcgaatgctactca 5942
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 GACCACTTGTGACAAACGCTTCATGCTTCTTACATTTAAACAGAGTCAATGTCACCTCA 181
Qy 5943 taagaagaagatcatcttaacacaaatgaagtgtctatltgagagacatgcagaaaagac 6002
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 CAAAGAAAGAGATTAATCTTAACCAATTTGAAGTTGCTATGCAAGACATGCGAAGAAAGAC 241
Qy 6003 acaggaagtgtcatttgaacacatcaggaatcccgcaagcccaaatgctcagatgt 6062
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 242 TCAGGAATTTGCAATTTGCAACATCAGAGATCAGAGACCCCAAAATGCTTCAAAATGCT 301
Qy 6063 acctcaaggaatcgttaagccacacagatgaatcaggggcctttggaagttgccaggttt 6122
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 302 CCTCCAGGCGTCTGTAGGAACAACACTGTAAACAGAGACCTCTGAGAGTTGCCAGGCTT 361
Qy 6123 tctgtctgaaataccctagtgaccacaaagctcttcagacatcataataaactgcgactctg 6182
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 362 CCTGTCTGAAATACCTGGTGAACCCGAGGCTTTCAGACATCAACAAGCTGGCGCTTTG 421
Qy 6183 ctttaagaatttctaagaagtgtgagaatgaccttaagaaaaaataagagcttaattg 6242
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 422 TTTCAAGGACTTCTTAAAGAGTGTGAGATGCTTACGAAAAAATGAAGCTTAATTTG 481
Qy 6243 gccggttcaaaaggagatcaaaagggaattg--ggaaacatctctgccttcaaaaggagc 6300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 482 ACCAGATCAAAAGAGATATCAAAAGAACTGAGAGAAATTAACATGCGCTTAAAGAGCG 541
Qy 6301 cctacagccctcagatca--cagaagtcctcaggtatccaagccagtatctgtcc 6358
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 542 TCTGCACCCACTGTATTAACAGAAAAATCCCTCAGCTATACAAAGCTGTGTGCTGTAC 601
Qy 6359 ctgcacagagatccttcagtcgaatgagcttcgcgaataatgatactctaactgaatg 6418
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 602 TTGCGACAGAGATTCCTTCAGCCGAATGAGCTTCCGAAAAATGGAACCTTAAT--ACA 659
Qy 6419 caattgtttatcatctgcaaaagccatglatccaacatcgagtggtgaagaatcat 6478
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 660 CACTTGTTCATTCATTTGAAAAAGAACCATGTATTCAACACTGAGGTG--AGGCTCAT 716
Qy 6479 tggaaaccaca--tggaaatggaatctcggaaatatatcatcttgaagaatcagatggcc 6537
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 717 TGAATAATTAAGACTTGGATGTCAATTCGAAAGCAATAGT--ATTGATTACTCTCTCGAAG 774
Qy 6538 aagaanaatcaaatgtagatgtttaacgcttgagaatcagtgctatggttttaattgt 6597
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 775 AATGTGCTCCAGATGAGGCTGTT-ACGTGTAGAGCTTTGGC-GTGGTTTTTAACCTT 832
Qy 6598 cgggttaacaagctgttactt---ttaagacatttaatgactcaaaagtacatacat 6653
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 833 CTGGTAACATGCTGTACTTTAAAAACAAGGTTTAAATGATTCAAGGTACAGTACAC 892
Qy 6654 attacacattattatcacatagctgaagttaaaattatcttaacttaagtctgattt 6713

```

```

Db 893 ATTTCATATTATTATACCTACCTAATGTT-AAATTATTTACTTTAAGTCTTGT 951
Qy 6714 tttaattatcacatctatagatca----tttggaccatttcaatgtagtaa 6768
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 952 TTTAATTATATATACCATTTATAGTCCATTCATTTTGTGAACCATTTAAATGTAGTAA 1011
Qy 6769 tgccttatttaaggtactcaaaaataatgtaatgctttac 6808
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1012 TGCTTATTTTAAAGGTACTATTAAATATGTGAATGTTTAC 1051

```

RESULT 9
 BG247108 931 bp mRNA linear EST 13-FEB-2001
 LOCUS BG247108
 DEFINITION 602359776F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488235 5',
 mRNA sequence.
 ACCESSION BG247108
 VERSION BG247108.1 GI:12756923
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE NIH-MGC http://mgi.nci.nih.gov/
 TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
 AUTHORS Unpublished (1999)
 JOURNAL
 COMMENT Email: cgapbs-r@mail.nih.gov
 Contact: Robert Strausberg, Ph.D.
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/MLM at:
 http://image.llnl.gov
 Plate: LLM10334 row: 9 column: 20
 High quality sequence stop: 728.

FEATURES
 source
 1..931
 location/Qualifiers
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4488235"
 /clone_lib="NCI_CGAP_Mam1"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="10 months, virgin"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; site.1: SalI;
 site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

BASE COUNT 275 a 198 c 228 g 230 t
 ORIGIN

Query Match 8.9%; Score 610.4; DB 10; Length 931;
 Best Local Similarity 90.3%; Pred. No. 4e-130;
 Matches 688; Conservative 0; Mismatches 66; Indels 8; Gaps 3;

```

Qy 5244 ctcaagatgtgtgatatccagatgaagaaggtatcgtccttggaataacttactga 5303
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 CTCACATGATGTGATATCTCCAGATGAAGAGCGCTTGTCTCGAAGTACTTCACTGA 60
Qy 5304 gtcaagacttgygattactggaacaagcagctgctcctctctatgctggaatgta 5363
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 ATCGGAGCTGGTGGCTTACTGGAACAAGCGGCTGCTTCCATTCAGGCGGCGCATGTA 120
Qy 5364 tgaagcagtttaagttcaaaagtactatctcctatcttaagaagtaatcgagatgc 5423
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 CGAAGCAGTTATAGGTTTCAAAAGTACTCTTCTATTCATGAACTAATCGGAGTGC 180

```


QY	5424	aaagaacactccacaattcattgtaacttcaagaagatccagaanaattggttcacaa	5483
Db	181	AAAAAGACCTCTCCACAACTTATG6CAAACTTCAGGAGGATTCAGCAAAATGTCATCA	240
QY	5484	gagtaacgtgcgtggaagagatgctgttgcaacctattcgttggtgtttatagaaacaa	5543
Db	241	G-----GATGCTAAGCGGATGTTTGGCACACTATTTCCGCTGTTGTTTATAGAAACCA	294
QY	5544	gttcgggatttggatlgaaacaagaatttgtttcaagaagacctgcaataaccaacttgc	5603
Db	295	GTTGCGGATTTGGATGATMACAGGAATTTGTTTCAAGAGGACCTGCAATTAACCAAACTCCG	354
QY	5604	agagatatctccacagatttggaggattttccgcgaagaagatttggaggagatggtgttga	5663
Db	355	AGAGATTTCTCACGATTTGGAGGATTTTATGAGAAAGATTTTGGAGAGATGTGCTTGA	414
QY	5664	agtaataaagaacctatctccttgaacaagttaataatgatctccaaagaagcatat	5723
Db	415	AGTATTCAGGACCTCTAATCTGTAGCAAGTGCAAATTTGATCCAAACAGGCAATAT	474
QY	5724	tcaagttacctatgttggagccatactttgacacataltgatlgaaagagaatacccta	5783
Db	475	TCAGATTACCTACGTGGAGGCCCTTTTGACACCTATGAGATGAAGACAGATCACTTA	534
QY	5784	tttcgacaaaataacatactctgttgatctcatgtactgtcaacccttactttagatg	5843
Db	535	TTT-GACAAAATTAATATCTCCGGCCTTTTCATGATCTGACACCCCTTCACTTTAATGG	593
QY	5844	ccgtgccaatgggaactctatgaacattccaagaagaagaccattctgaactga	5903
Db	594	CCGTGCTCATG6GAGCTCATGACACATTCAAACGGAAGACCATTTCTGCACACGTCTCA	653
QY	5904	tg-ctttccctatataaacaagggtcaatgtcaactcaataaagaagatcatcttaa	5962
Db	654	TGTCTTTCTTACATTAATAACAGAGATTCATGTCATCTCAAAAGAAATATCTTAA	713
QY	5963	caccaattgaagtgtctatttgaaggacatgcagaaagaagacac	6004
Db	714	CACCAATTGAAGTGCTATCCGAAGACTGCAGAAACAGACC	755
RESULT	10		
LOCUS	BE891826	958 bp	linear
DEFINITION	60143427F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3919720 5',		EST 20-OCT-2000
ACCESSION	BE891826		
VERSION	BE891826.1	GI:10351537	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;		
AUTHORS	Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo.		
TITLE	1 (bases 1 to 958)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: ATCC/DCTM/DMP		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: L1AM9749 row: O column: 17		
	High quality sequence stop: 637.		
FEATURES	Location/Qualifiers		
source	1..958		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:3919720"		

Query Match	Best Local Similarity	8.8%;	Score 599.4;	DB 10;	Length 958;
Matches 742;	Conservative	0;	Mismatches 101;	Indels 22;	Gaps 6;
QY 4887	ggaattgagggaaacaacatttcctgatccaggltccaggatctcgtgtttlcaatctccat				
Db 1	GGAAATGAGCGAACAACATTTCTGATCAGTCCAGATCTGTTTCAATCTCATAT				
QY 4947	gattctttctgatactcgttgaataatgaagaacaaccagggaggttcctgaaatttgatga				
Db 61	GATTCCTTTCTGATCTGTAATAATGAAGGAAACACAGGAGGATCCGAATGTGATTTGA				
QY 5007	tctaatgtacagaatctgacgaagggtttacagaccctccaga---gcgaatgacctggt				
Db 121	TCTAATGTACAGAAATTTGCCAAGGTTTACAGACCTTCCAGATCTGCATTTGACCTGGTT				
QY 5064	gcagaacatgycagagcaagcactcagaaagcaalcalgtcgaagctgcagcagtgct				
Db 181	GCAGAACATGGCAGCGACACCTCAGAACGAAAGCATTCATGCTGAAAGCTGCACAGCTCT				
QY 5124	agtcacccaagcaagcaactgtgttcgtgaataatttgagcagctcgtggagaccggaatatct				
Db 241	AGTCCACTCAGCAGCCTTGTGCTGAATATTGACCATCTGGAGACCGGAAATATCT				
QY 5184	tctcgtggagatggtacacattccagaatttcaatctcaatgatttttagaagaatctcggt				
Db 301	TCTCTGGAGATGTTGATACATTTTGAGATATTTCATCTTAATGTTTGAAGAATCTCGGT				
QY 5244	ctcagaatgltgttalcctccagatbaaagaaggtatctcgtcctgaaaaacttacttcga				
Db 361	CTCAGATGATGTGTGATCTCTCCAGATGAAGAAGGTATCTGCTGGAAATATTCTTACTGA				
QY 5304	gtcagaactgttgagattactggaacaagcagctgcttcctctcctcgtcgtgcatgtga				
Db 421	GTCAGACTTGTGAGATTACTGTGMAACGACGCTGCTCTCTCTATGAGCTGGCATGTA				
QY 5364	tgaagcagttaatgaagatttacaagaattactat-tcctattcagaagcctcaatcggagtg				
Db 481	TGAGAGATTAAATGAAGTTTACAAAGTACTTATCTCTATTCATGAAAGCTTAATCGGGATG				
QY 5423	caagaagaactcacacaatcattcattgctaaacttcaagaagaagcattcagaacaattgtcaltc				
Db 541	CAAGAACAATCATCCACATTTCAATGATGAACCTTCAAGAGATTTACACAAATTTGTTTCATC				
QY 5483	agaagctcgtcgtggagcggatgttggcaactatttcgttgttggtt--ttaatgaac				
Db 601	AGGATGTGTAC-----CGGATGTTTGCCACCTAATTTTCTGTGGGTTTCTATGTAAC				
QY 5541	caagttcgg-----gatttgtatgaacaagaattgattgtttacaagaagcctgcaat				
Db 655	CCAAGTCCGGGGCAATTTGGGCTTGAACCCGCAATTTGTTTCCAGGGAGGCTTGGCCATA				
QY 5592	aaccaaacctgcagagatactcac--agattgagagattttagcggagaagaatttgag				
Db 715	ACCCAAACTTGGCGAGATTTCTCACAAGAAATGGGAGGATTTACGCCGAGAAAGACTTGGCC				
QY 5651	aggatgtgttgtgaagtaatcaaaagactctaactcgttagacaagtgtlaattgatccta				
Db 775	AAGCATGTGTGTGAGCTATACAAAGAACTATGCTGTACAGTTATTAACTCAGGCTT				
QY 5711	acaagcattatattcagattactcta				

Db 835 ATTACTCATGTTGCCATCTCCTA 859

RESULT 11
AA429436 588 bp mRNA linear EST 16-OCT-1997
LOCUS zw47a12.f1 Soares-total_fetus_Nb2HP8_9w Homo sapiens cDNA clone
DEFINITION IMAGE:773182.5' similar to WP:F46H5.4 CE04590 RAT TRG GENE PRODUCT
; mRNA sequence.

ACCESSION AA429436
VERSION AA429436.1 GI:2112516
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 588)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geiseli, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,
Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie,
'T', Waterston, R., and Wilson, R.
WashU-Merck EST Project 1997
Unpublished (1997)
CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28ml3 rev2 ET from AmerSham
High quality sequence stop: 495.
Location/Qualifiers
1. 588
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:773182"
/clone_lib="Soares-total_fetus_Nb2HP8_9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dt) primer [5'
TGTACCAATCTGAGTGGAGCGCCGCTTATATTTTCTTTTCTTTT 3']
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 202 a 99 c 124 g 163 t

ORIGIN

Query Match 8.6%; Score 588; DB 9; Length 588;
Best Local Similarity 100.0%; Pred. No. 5.2e-125;
Matches 588; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 accagatgaagaagaacataatagaaggctagtgtctctgatatatcccaagaacattt 803
|||||
Db 1 ACCGATGAGAGAAACCAATAGACGGCTTACTGTCTCGATATACCCAAAGACATT 60
|||||

QY 804 tggctcaagaagctctgttaaatgctatcacatcaagtttgaattgaattgaaccat 863
|||||
Db 61 TGGTCAAAAGACTTCTTTAAATGCTTATATCACTCAAGTTTGAATTTGAATTTGAACCAAT 120
|||||

QY 864 ttgttcaagtttggcttatatgagtgcaaggaagaagaatttcagaaacttta 923
|||||
Db 121 TTTTGCAGAGTTTGGCTTATATGATGTCAAGAAAGAAAGATTTCAGAAAACTTTTA 180
|||||

QY 924 ttgttgaactaatcttgagagatgaagaaggtgttgtaagtcacatgataccacgtctgc 983
|||||

Db 181 TTTTGACCTTATTCTGAGCAGATGAAGGCTTGTACGTCCACATGTACACCTCTCC 240

QY 984 cattaactaccctggcaagatcagcaatltttctatcaattatccttcccaagatgttt 1043
|||||

Db 241 CATTACTACCTTGGCAAGATATCAACAATTTTCTATCACTTATCTCTCCAGAGATGTTT 300
|||||

QY 1044 tctgttaataaagctgaagaagctccctcagcagaaggaagcatttggagatgtgcagaac 1103
|||||

Db 301 TCTTTGATTAAGCTGAAAGAGTCTTACAGAGAGACATTGGAGAGTGTGCAGAAC 360
|||||

QY 1104 atataatgatttcaagaagcagatgccaccaagaataagaanaaacttgaaactgaa 1163
|||||

Db 361 ATATATGATTTTCAAGAGAGATGCCACCAAGATTAAGAAAACTGAGAAACTGAA 420
|||||

QY 1164 gagtcagaagatcagtttgcgaagacttgggaatatgcatagcctttgttggac 1223
|||||

Db 421 GAGTCAAGCAGATCACTTTGGCCAAAGACTTGGAAATATCCATGCTTTGCTTGAC 480
|||||

QY 1224 ttgaatccatttaataatattgttagcagtgctctggagtttggaagaagatctacaga 1283
|||||

Db 481 TGCATTCATTATGATATATTTGTTAGCAGTCTCGAGATTGGAAAGAGATTTCACGA 540
|||||

QY 1284 agtagaatcagctactgagaaacgaagaaggtctgtgtcagagagag 1331
|||||

Db 541 AGTAGAATCATGACTGAGAAAGCAAGAGCTTGTGTACAGAGAG 588
|||||

RESULT 12
BC020473/C 1510 bp mRNA linear HTC 03-JAN-2002
LOCUS BC020473
DEFINITION Homo sapiens, clone IMAGE:3853958, mRNA.
ACCESSION BC020473
VERSION BC020473.1 GI:18042973
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1510)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
CONTACT: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalobcm.tmc.edu
Villalob, D.K., Luna, R.A., Hale, S.M., Huliy, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gunatone, P., Yoon, V., Kowis, C., Martin, R.,
Lawrence, S., Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAC Plate: 13 Row: B Column: 12
This clone has the following problem: no polyA-tail.
Location/Qualifiers
1. 1510
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3853958"
/tissue_type="Colon, adenocarcinoma"
/clone_lib="NIH_MGC_65"
/lab_host="DH10B"

BASE COUNT 287 a 431 c 456 g 336 t
 ORIGIN

Query Match 8.5% Score 580.6; DB 11; Length 1510;
 Best Local Similarity 66.9%; Pred. No. 3.6e-123;
 Matches 858; Conservative 0; Mismatches 419; Indels 6; Gaps 2;

5132 cagcagcctgtgtgtaaatatttgagcatgctgagagaccggaatatcttctgtg 5191
 1509 CGGCGCGCTGCTGAGTACCTCGCTGCTGAGAGACCGCCACTGCGCTGG 1450
 5192 gatgttaacattcaagatatcttaacatgatttlaagaaatctgcggtcagatg 5251
 1449 GCTGGCTTCTTCAGAACATCTCATTCATGCTAGAGAGATCCGCCATCTCCGACG 1390
 5252 atgtgtatctccagatgaagaaggtatctgctctgaaatacttacttgaaggaac 5311
 1389 ACATCTGCTCCGACGAGAGAGGAGGCTTCTGCTCCGGAAACACTTACTGAGCTGGGCG 1330
 5312 ttgtggattctggaagaagcagctgtctctctctatgctgctgcatgtatgaagcag 5371
 1339 TGTATGGGTCTGTGAAAGAGAGCGGCTACTTACCATGGCGGCTCTACAGAGCGG 1270
 5372 ttaatgaagttacaagaacttctctatctatgaagcaatcagagatgcaagaagac 5431
 1269 TGAATGAGGTCTACAGAAAGCTCATCCCATCTCTGGAAGCCCGGATGATACAGAAAGC 1210
 5432 tatccacaatcattgttaaacctcaagaagatccagaagaattgttcaatcagaagctg 5491
 1209 TGGCGCGGTGCGACGAGCAACTGAGAGGCGCTTACCATGATGATGACCAAGATTCCG 1150
 5492 gctggagacgagatgttggacacctatttctgtgtgtttatggaaacagttggg 5551
 1149 GCTGGAGACGGGTGTCGGAGATATTCGCGTGGGCTTACGGCGCCACTCTGGTT 1090
 5552 atttgaagaacaagaattcttcaagagcgctgcataacaagaacttcaagatat 5611
 1089 GCGTGGATGAGAGAGAGTTTGTGTACAGAGAGCCATGATCAGAGAGTGGCAGAGATCT 1030
 5612 ctccagaattggaagatttcaaggaagaatttggagaagatgtgttgaatgaatca 5671
 1029 CACACGGGTGAGGAGGATTTACAGGAGATTTGGCGAGAGCTCTTATGATTAATCA 970
 5672 aagactctaatctctgtaagaagtgtaaatgaatccctaaagaagcatatgaatga 5731
 969 AAGACTCTAACCTCTGACAGTCCAGCTTGAATCAAAAAGGCTTACATCCAGATCA 910
 5732 cctatgtgagacatacttgcacacatatgatgaagaagaacacacatttgcaga 5791
 909 CCTATGTGGAACCGTACTTTGATACCTAGAGCTCAAGACCGGGTGACACTTTGAC 850
 5792 aaatataacatcttcgtgatcatctgtaactgtaacaccttactttagatggcggtgcc 5851
 849 GCACCTATGGGCTTGACACATCTCTGCTGACCCGCTTACGCGGATGGGCGCGCAC 790
 5852 atgggaacttcatataaattcaaaagaagacacattctgactacgctcatgcttcc 5911
 789 AGGGGAGCTGCGCGAGACACAGAGCGTAAAGCCTGCTGACACCGACACAGCGCTTCC 730
 5912 ctatataataaagaaggttcaatgtcactcatlaagaagaagatcatcttaacacacattg 5971
 729 CCTACATCAAGACCTGCATCCGTGTGTGACACGGGAGAGACGGTGTGACACCGAGTGG 670
 5972 aagttgcatatgagacatgacagaaagaacagagattgtgcatcttgcacacacag 6031
 669 AGGTGGCCATTCAGAGACATGAGAAAGACACGGGAGCTGCGCTTTTCCACCGCAGG 610
 6032 atcccgcaaccaccaaatatgctcagatgtaatcaccaggaatctgtaggcacacagctga 6091
 609 ACCCACAGATGCTAAGATGCTACAGATGCTTCAAGGCTCTGTAGGGCCACCGATGA 550

6092 atcaaggagccttggagaagtgtcccaaggtttctctgtcgtgaatlaacctgagcccaagc 6151
 549 ACCAGGGTCCCTGGAGGTGGCCAGGTGTTTATGACAGATGCCGGAAGACCCCAAGC 490
 6152 tcttcaagacatcataaactgcgactctgtctttaaagatttctaaaggtgtgag 6211
 489 TCTTCCGGATCAACAATAATTCGGCTCTGCTTCAAGACTTGTGCAAGAAATGTGAGG 430
 6212 atcccttaagaanaaagaagcttattggccggttcaagaaggaatlaacaaaggaat 6271
 429 ATCGCGTGGCGGAAATATMAGCCCTTATTTGGCCGAGACAGAGATACACCGCTGAGC 370
 6272 tg--gggaactatctctgccttaagaagcgccctacagccctagatcaagaagaatccct 6329
 369 TGGAGGCAACTACTCTCGCTGCGGAGGCTCTGACAGCCCTGCTTACCGAGCCCTGTC 310
 6330 cagttatccagccagatattgtctgcccctgcacaaagatcttctcagtgatgagc 6389
 309 CCCAGCTGATGACCCACCC--ACCCGGCTCAGAGACTCTTGAACAGACCAAGT 254
 6390 ttgcgaatgagatctcaaac 6412
 253 TTCCGAAGGACAGACTCTGAGC 231

RESULT 13

B1689126 869 bp mRNA linear EST 18-SEP-2001
 LOCUS 603315930F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5355798 5',

DEFINITION mRNA sequence.
 ACCESSION B1689126
 VERSION B1689126.1 GI:15651755

KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 869)
 AUTHORS NIH-MGC
 TITLE NIH-MGC http://imgc.nci.nih.gov/.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgabds-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM11904 row: 1 column: 07
 High quality sequence stop: 761.

FEATURES

source

1..869
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5355798"
 /clone_1lb="NCI_CGAP_Mam6"
 /sex="Female, Virgin"
 /tissue_type="Infiltrating ductal carcinoma"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site:1; Salt: 1.0M; pH: 8.0; Temp: 65°C; Method: PCR; Library constructed by Life Technologies, Inc. Investigator providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT

223 a 213 c 237 g 196 t

Query Match 8.4% Score 572; DB 10; Length 869;
 Best Local Similarity 86.7%; Pred. No. 3e-121;

OM of: US-09-737-246-2 to: EST:* out_format : pfs
Date: Jul 16, 2002 1:13 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+2.pn.model -DEV=xih
-O=Cg92.1/USFTO.spool/US09737246/runat.16072002.075429.26451/app-query.fasta.1.2180
-DB=EST -QPM=fastap -SUFFIX=p2n.rst -GAP=12.000 -GAPEXP=4.000
-MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-OGAPEXT=0.050 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-LIST=45 -DOCALLIC=200 -MATRIX=blonum62 -TRANS=human40.cdi
-ALIGN=15 -MODE=LOCAL -OUTPMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09737246_ECGN1.1.4507
-NCPU=6 -ICPU=3 -LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30
-NO_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-09-737-246-2

Query length: 2090

Database: EST*

Database sequences: 13736207

Database length: 1841457050

Search time (sec): 3201.680000

score_list:

Sequence	Strd Orig	ZScore	EScore	len	Documentation
gb_hic:AK016495	+ 3187.00	5492.05	1.2e-296	2710	AK016495 Mus musculus adult ma
gb_hic:BC018599	+ 2155.50	3709.13	2.4e-197	1311	BC018599 Mus musculus, clone 1
gb_hic:AK004816	+ 1821.50	3119.20	1.7e-164	2694	AK004816 Mus musculus adult ma
gb_hic:BC020473	+ 1659.00	2844.46	3.4e-145	1510	BC020473 Homo sapiens, clone 1
gb_hic:AK010755	+ 1396.50	2388.32	8.8e-124	1507	AK010755 Mus musculus ES cells
gb_est1:AJ397707	+ 1293.50	2217.76	2.8e-114	796	AJ397707 AJ397707 dkfz426 Gallu
gb_hic:AK018051	+ 1219.00	2078.11	1.7e-106	1712	AK018051 Mus musculus adult ma
gb_est2:BG247108	+ 1209.50	2069.71	4.9e-106	931	BG247108 602359776F1 NCI CGAP
gb_hic:BC018075	+ 1199.00	2042.23	1.7e-104	1873	BC018075 Homo sapiens, clone 1
gb_hic:BC018076	+ 1199.00	2042.23	1.7e-104	1873	BC018076 Homo sapiens, clone 1
gb_est2:BI688126	+ 1193.00	2041.95	1.7e-104	869	BI688126 60331530F1 NCI CGAP
gb_est2:BG872239	+ 1146.00	1962.31	4.7e-100	795	BG872239 602790848F1 NCI CGAP
gb_hic:AK013336	+ 1106.00	1894.89	2.7e-96	636	AK013336 Mus musculus 10, 11 de
gb_est2:BI770146	+ 1104.00	1888.45	6.1e-96	796	BI770146 603053332F1 NIH_MGC_12
gb_est2:BF309712	+ 1092.00	1866.00	1.1e-94	898	BF309712 601891854F1 NIH_MGC_17
gb_est1:AM918177	+ 1088.00	1863.61	1.1e-94	636	AM918177 EST350481 Rat gene ind
gb_est2:BI088758	+ 1085.00	1855.45	4.2e-94	795	BI088758 602851547F1 NIH_MGC_10
gb_est2:BI101016	+ 1076.00	1841.91	4.2e-93	678	BI101016 602846468F1 NCI CGAP
gb_est2:BB891826	+ 1072.00	1830.39	1.0e-92	958	BB891826 601434272F1 NIH_MGC_72
gb_est2:BB717504	+ 1029.00	1760.27	8.4e-89	676	BB717504 f445f10.Y1 zebrafish v
gb_est2:BF159615	+ 1027.50	1753.38	2.0e-88	935	BF159615 601767221F1 NCI CGAP
gb_est2:BI737331	+ 1024.50	1753.50	2.0e-88	927	BI737331 603357985F1 NIH_MGC_94
gb_est2:BI820302	+ 1024.50	1748.28	3.9e-88	927	BI820302 603035786F1 NIH_MGC_11
gb_est2:BB531136	+ 1022.00	1745.80	5.4e-88	805	BB531136 601278315F1 NIH_MGC_39
gb_est2:BI078974	+ 1001.00	1709.78	5.5e-86	588	BI078974 602873230F1 NCI CGAP
gb_est1:AA424035	+ 999.00	1709.98	8.1e-85	652	AA424035 zw47a12.Y1 Soares tota
gb_est1:AM742035	+ 988.00	1680.61	2.3e-84	652	AM742035 u52c04.Y1 Soares, mous
gb_est2:BI733553	+ 983.00	1680.61	2.3e-84	652	BI733553 603355485F1 NIH_MGC_94
gb_est2:BB571095	+ 974.50	1659.39	3.5e-83	1079	BB571095 601334443F1 NCI CGAP
gb_est2:BI159195	+ 973.50	1661.97	2.5e-83	778	BI159195 602922704F1 NIH CGAP
gb_est2:BB677746	+ 968.00	1650.73	1.1e-82	884	BB677746 602625421F1 NCI CGAP
gb_est1:BB677746	+ 967.50	1653.48	7.1e-82	672	BB677746 602625421F1 NCI CGAP
gb_est1:BB677746	+ 959.50	1636.67	6.5e-82	842	BB677746 603381455F1 NIH CGAP
gb_est2:BF141712	+ 954.50	1629.67	1.4e-81	737	BF141712 60179970722F1 NCI CGAP
gb_est2:BF120001	+ 954.00	1630.50	1.4e-81	648	BF120001 601758086F1 NCI CGAP
gb_est2:BB857715	+ 943.00	1628.64	1.8e-81	654	BB857715 7496903.X1 NCI CGAP
gb_est1:AA275395	+ 942.00	1611.41	6.0e-80	567	AA275395 va92f11.Y1 Soares mous
gb_est2:BB678867	+ 929.00	1583.36	6.0e-79	857	BB678867 602624767F1 NCI CGAP
gb_est2:BM310731	+ 926.00	1583.42	6.0e-79	575	BM310731 f447g06.Y1 H885 islet
gb_est2:BB698863	+ 924.00	1580.77	8.4e-79	540	BB698863 RC4-NN0025-120600-016

gb_est2:BG080811	+ 913.00	1560.38	1.1e-77	595	BG080811 H3057607-5 NIA Mous
gb_est2:BG927424	+ 906.00	1547.67	5.9e-77	620	BG927424 HNC53-1-B8.R HNC (H
gb_est2:BF036757	+ 903.00	1542.76	1.1e-76	606	BF036757 601459988F1 NIH_MGC
gb_est1:AI653716	+ 896.50	1530.30	5.4e-76	662	AI653716 wb36e07.X1 NCI CGAP
gb_hic:AK014226	+ 884.00	1502.47	1.9e-74	1051	AK014226 Mus musculus 13 da

seq_name: gb_hic:AK016495

seq_documentation_block:

LOCUS AK016495 2710 bp mRNA linear HTC 19-JAN-2002
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4931431C02:homolog to KIAA1395 PROTEIN (FRAGMENT), full insert sequence.

12 GGAAGCACCATCAACCCATCTGTGGCCATGGCCATCGCTGGGG... 55

[illegible]


```

|||||.....|
922 GTCGCTCTGGGGACACTCAGAACTTAGGAGAGCATTTGAGAA 971
1609 rgserleuylsthrilleuthrlyragluaglulaspheugluuarg 1625
|||||.....|
972 AGTCCCTCAGACCATCTCTGACATACGAGAGGAGCATATGGCTGAGG 1021
1626 gluThrThrPheProaspGluValGlnaspLeuValPheasnLeuHisMe 1642
|||||.....|
1022 GACAGACACCTTGCTGACAGAGTCCGAGACTCATGTTCAACCTGCACAT 1071
1642 ttleuSeraspThrValylsMetlysGluHisglnluaspProglu 1659
|||||.....|
1072 GATCCTGACAGACACGCGTAAGATGAAGAACACAGAGAGACCTGTAGA 1121
1659 etleuIleaspLeuMetlyrThrIlealysGlyTyrGlnThrSerPro 1675
|||||.....|
1122 TGCTCATGGACCTCATGTACAGAAATGCTCGGGATACCAAGGCTCTCCA 1171
1676 glu...ArgleuThrTripleuglnaspMetIleagllylshisSerGluAr 1691
|||||.....|
1172 GACCTCGCGCTGACATGGCTGCAGAACATGGCTGCACAAACATGCAGAGCT 1221
1691 gSerasnHisalaglualalaglncysleuValHisseralaleuVal 1708
|||||.....|
1222 GGGCAATCATGCAGAGGCTGCCAGTCATGATGCACCGCGCTGCGCTGG 1271
1708 alalaglulThrleuSerMetleuGluaspArglylserTyrleuProvalgly 1724
|||||.....|
1272 TGCGCGAATACCTTCCTCCTCGAGGACAGTCCGACCTGCTGTGGGC 1321
1725 CysValThrPheGlnaspIleSerSeraspValleugluGlnSerAlaVal 1741
|||||.....|
1322 TGCTGTCTCTCCAGAAATGTCATCCAAAGCTGCGAGAGATGTCGCAT 1371
1741 lseraspaspValValSerProaspGluGlyllecysSerGlylser 1758
|||||.....|
1372 CTCTGATGATCATCTGCGCCAGATGAGAGGCTTCTGCTTAGAAGA 1421
1758 yrrPheThrGluSerGlyleuValIleGlyleuGlnGlnAlaIleaser 1774
|||||.....|
1422 ACTTCACAGAACTGGGTGCTGGGGCTGCTGAGCAGGCGCGCTAC 1471
1775 PheSerMetIleagllyMetlyrGluValalaspGluValTyrlyValle 1791
|||||.....|
1472 TTCACCATGGGTGCTGTATGAAAGCGGTGAACGAAGTCTACAAAACCT 1521
1791 uilleProIleHisGluAlaAsnArgaspAlalysLysLeuSerThrIleH 1808
|||||.....|
1522 TATCCCATCTCTTGAAGCCACAGACATACAAAGAGCTGGCTGGGTGC 1571
1808 lsglyllysleuGlnGluAlaPheSerlylIleValHisGlnSerThrGly 1824
|||||.....|
1572 ACGGGAACTGCAGAGAGCGCTTCCACCAAGATTATGCACAGAGCTGTGC 1621
1825 TrpGluArgMetPheGlyThrTyrPheArgValIleGlyPheTyrGlyThrly 1841
|||||.....|
1622 TGGGAGCGGTATTTGGACATATTTCCAGAGGGCTTCTATGGACACAG 1671
1841 spheGlyaspLeuaspGluGlnGluPheValTyrlysgluProalilet 1858
|||||.....|
1672 ATTGGTGCAGCTGATGACAAAGAGTTTGTACAAAGAACCGTCAATCA 1721
1858 hrllysleuIleaglulIleSerHisArgleuGlnGlyPheTyrGlyThrly 1874
|||||.....|
1722 CGAAGCTTCAGAGATCTCACACCGGCTGAGAGTCTATACGGAAGG 1771
1875 PheGlyGluaspValValGluValIlelysaspSeraspProValasp 1891
|||||.....|
1772 TTGGGGAGATGATGTGTAGACATCAAAAGTTCACACCGAGTGCAGCA 1821
1891 scylsLysleuaspProaspLysAlaTyrIleGlnIleThrTyrValGluP 1908
|||||.....|

```

```

1822 GTCGAAGCTGGACCCACAGAGGCGTACATACAGATTAACCTATGTGGAGC 1871
1908 rotyrPheaspThrThrlyGluMetLysaspArgIleThrTyrPheaspLys 1924
1872 CACATTTGCAGACTTATGAGCTCAAGATCGGGTGCACCTACTTTCATTCGG 1921
1925 AsnTyrAsnLeuValrgrPheMetTyrCysThrProPheThrLeuaspGlu 1941
1922 AACTATGGCGTGGGGCGCTTCCTCTCTGCACACCTTGCACACCGAGATGG 1971
1941 yarGalalHisGlyGluLeuHisGlnGlnPheLysArglylserThrIleuThr 1958
1972 ACCTGGCGACAGGAGAGTTGGCCGACACACAAACCGACAGAGCTGTCTGA 2021
1958 hrlhrSerHisalalPheProTyrIleLysThrArgValaspValThrHis 1974
2022 GCAGGGAGCATGCGCTTCCCTACATCAAGACAGCATCCGAGTGTCCAC 2071
1975 LysGluGlnIleIleLeuThrProIleGluValalalegluaspMetGlu 1991
2072 CGTGAGAGACAGATGCTGCACACAGTGAGAGTGGCCATTGGAGACATGCA 2121
1991 nLysLysThrGlnGluLeuAlaPhealThrHisGlnaspProalasp 2008
2122 GAAGAAACACCGGGAGCTGCGCTTGGCACCGAGCAGGAGACCTCCAGATG 2171
2008 rolysmetleuGlnmetValleuGlnGlyserValGlyThrThrValasn 2024
2172 CCAAGATGCTGCAGATGTTCTCTCAGAGGTTCTGTGGACCCACTGTGAAAC 2221
2025 GlnGlyProleuGluValalalaglinalPheLeuSerGluIleProSerasp 2041
2222 CAGGCTCCCTTGGAAAGTGGCCCAAGGTGTTTGTGCAGAGATCCCAAGA 2271
2041 proLysleuPheargHisHisasnLysleuArgleuCysPheLysasp 2058
2272 TCCCAACCTCTCCGACATCAACAAACCTCGGCTGCTGTTAAAGATT 2321
2058 heThrLysArgCysGluaspAlaleuArgLysasnLysSerleuIleGly 2074
2322 TCTCAAAAAGTGGAGAGATGCACCTGGAAGAAAGAACAGAGCCCTGATTGGC 2371
2075 ProValGlnLysGluTyrGlnArgGluLeu 2084
2372 CCAGACCAAGAAAGATGATCCACCGGGAGCTG 2401

seq_name: gb_hlc:BC018599
seq_documentation_block:
LOCUS BC018599 1311 bp mRNA linear HTC 03-DEC-2001
DEFINITION Mus musculus, clone IMAGE:3596745, mRNA.
ACCESSION BC018599
VERSION BC018599.1 GI:17235789
KEYWORDS HTC.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submision
JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing by: Baylor College of Medicine Human Genome

```

Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: villalobos@bcm.tmc.edu
 Villalobos, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
 Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 16 Row: e Column: 2
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis

This clone has the following problem: no 5' EST match.

Location/Qualifiers

1. 1311
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:3596745"
 /tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
 ductal carcinoma, 5 month old virgin mouse."
 /clone_1lb="MCI_CGAP_Mam6"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORE6"

BASE COUNT 363 a 309 c 333 g 306 t
 ORIGIN

alignment_scores:

Quality: 2155.50 Length: 437
 Ratio: 4.978 Gaps: 1
 Percent Similarity: 99.085 Percent Identity: 97.712

alignment_block:

US-09-737-246-2 x BC018599 ..

Align seg 1/1 to: BC018599 from: 1 to: 1311

```

1271 CysIleAlaThrAspAspTyrGluSerGluSerGlySerMetIleSerG 1287
||||| |||||||:|||||
1 TGTATAGCCCGCATGAGTATGACAGTGAAGTGGAGCATGATGAACCA 50
1287 nThValAlaMetAlaIleAlaGlyThrSerValProGluLeuThrArgp 1304
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
51 GACGGTTCGCAATGCGAATTCGAGAACATCCGCTCAGCTGACAAGAC 100
1304 roGlySerPheLeuLeuThrSerThrSerGlyArgGlnHisThrPhe 1320
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
101 CTGGCAGTTTCTCTCCTCAGTCAAGAGTGGCGGACACACACACCTTC 150
1321 SerAlaGluSerSerArgSerLeuLeuIleCysLeuLeuThrValLeu 1337
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
151 TCAGCAGAGTCACAGCAGAGCCTTTGATCGTCTGCTTGGGTTCCAA 200
1337 sAsAlaAspGluThrValLeuGlnLysTrpPheThrAspLeuSerVal 1354
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
201 GAATGCAAGATGAACGGTTCTGCAAAATGGTTACAGATCTCTCACTCC 250
1354 eugInLeuAsnArgLeuLeuAspLeuLeuThrLeuCysValSerCysPhe 1370
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
251 TCCAGCTGACACCGGTGTGTAGATCTGCTTACCTGTGTGTCTGCTTT 300
1371 GluTyrLysGlyLysValPheGluArgMetAsnSerLeuThrPhe 1387
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
301 GAGTACAAAGCAAAAGGATTTGAACGATGAATGATTAACTTTAA 350
1387 sLySerLysAspMetArgAlaLysLeuGlnGluAlaIleLeuLysSer 1404
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
351 GAAATCAAAAGACATGAGAGCCAGCTTGAAGAACCATCTCGGGAGCA 400
1404 legAlaIaArgGlnGluMetValArgArgSerArgGlyGlnLeuArg 1420
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
401 TCGGTCCAGGCAAGAAATGTTACGGCGAGCCGAGGGCAGCTCGAGAG 450

```

```

1421 SerProSerGlySerAlaPheGlySerGlnGlnAsnLeuArgTrpArg 1437
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
451 AGCCCATCTGGAAGCGCTTTGGAGGCAAGAAACCTGGCGGAGAA 500
1437 sAspMetThrHisTrpArgGlnAsnThrGluLysLeuAspLysSerArg 1454
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
501 GGACATACCTACTGGCGGTGAGAACTCAGAGAACTCGATTAAGTCAAG 550
1454 legLuiIleGlnHisGluAlaLeuIleAspGlyAsnLeuAlaThrGluAla 1470
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
551 CAGAGATAGAACAAGCAAGCACTGATTGATGAACCTGGCTACTGAAGA 600
1471 AsnLeuIleIleLeuAspThrLeuGluIleValValGlnThrValSer 1487
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
601 AACCTCATCATCTTAGACAGCTAGACATCATGCTCCACACTGTGTCT 650
1487 TrnGluSerLysGluSerIleLeuGlyValLeuLysValLeuLeu 1504
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
651 AACGGAATCCAAAGAGAGTATCTCGGTGTGTGTCTAAAGTCTGTAC 700
1504 IsSerMetAlaCysAsnGlnSerAlaValTyrLeuGlnHisCysPheAla 1520
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
701 AGAGCATGGCTGCAACCAAGAGTCACTATCTGACGACACTGCTCGCC 750
1521 ThrGlnArgAlaLeuValSerLysPheProGluLeuLeuPheGlnGlu 1537
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
751 ACCGAGAGACCTCGTCTCAAAAGTTTCCAGACTGTCTTTGAGGAGAA 800
1537 uThrGlnGlnCysAlaAspLeuCysLeuArgLeuLeuAlaGlnHisCys 1554
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
801 GACAGAGCAAGTGGCGATCTGTGCTCGCGGTTCCTCCCTCAATGTGAG 850
1554 eTrSerIleGlyThrIleArgSerHisProSerAlaSerLeuTyrLeu 1570
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
851 GCAGCATCACTACATCCGCTCGACGTAAGTCCCTTACCTACCTCC 900
1571 MetArgGlnAsnPheGluIleGlyAsnAsnPheAlaArgValLysMet 1587
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
901 ATAGGCAGACACTTGTAGATCGGGAATTAATTTGCCAGAGTGAATAAT 950
1587 nValProMetSerLeuSerSerLeuValGlyThrSerGlnAsnPheAsn 1604
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
951 AGTACAGATGTCACTGCTCTTGGTGGTGAAGTCTCAGAATTTTAAG 1000
1604 InGluPheLeuArgArgSerLeuLysThrIleLeuThrTyrAlaGlu 1620
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1001 AAGAATCTTAAACGCTCTTAAAAATATTGACATATGCTGAAGAA 1050
1621 AspLeuGluLeuArgGluThrThrPheProAspGlnValGlnAspLeu 1637
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1051 GATCTGAATGTGAGGAACACACATTTCTGATCAGGTCCAAAGATTTG 1100
1637 lPheAsnLeuHisMetIleLeuSerAspThrValLysMetLysGlnHis 1654
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1101 CTTCATATCTCATATATGATCTTTCTGACACGTGAAGATGAAGAGCA 1150
1654 InGluAspProGluMetLeuIleAspLeuMetTyrArgIleAlaLysGly 1670
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1151 AGGAAGACCCAGAAATGATGATTGATCATGATAGATGCAAGAGGC 1200
1671 TyrGlnThrSerProGlu...ArgLeuThrTrpLeuGlnAsnMetAla 1686
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1201 TACACACAGCTTCACAGACCTGCGCTGACCTGTGTCAGAACATGGCT 1250
1686 yLysHisSerGluArgSerAsnHisAlaGluAlaAlaGlnCysLeuVal 1703
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1251 CAAGCACTGTGAAGCAAGCAATCATGCGCAAGCTGCTGAGTCCGTG 1300
1703 IsSerAlaAla 1706
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1301 ACTCCGCAAGCC 1311

```

seq_name: gb_hic:AK004816

seq_documentation block:

LOCUS AK004816 2694 bp mRNA linear HTC 19-JAN-2002

DEFINITION Mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200017A24:homolog to BA1565F24.1.1 (NOVEL PROTEIN SIMILAR TO DROSOPHILA CG6630 AND CG11376, KIAA1058, RAT TRG (ISOFORM 1)) (FRAGMENT), full insert sequence.

ACCESSION AK004816

VERSION AK004816.1 GI:12836282

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (strain:C57BL/6J) adult male lung cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library clone:1200017A24.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (sites)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

JOURNAL MEDLINE PUBMED 99279253 10349636

REFERENCE 2 (sites)
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL MEDLINE PUBMED 20499374 11042159

REFERENCE 3 (sites)
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Iwawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,T., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL MEDLINE PUBMED 20530913 11076861

REFERENCE 4 (sites)
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

JOURNAL MEDLINE PUBMED 11076861

REFERENCE 5 (bases 1 to 2694)
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arkawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fudusada,S., Fukunishi,Y., Furuno,M., Hanganaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hirooka,T., Hori,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Ouackembush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schirral,L., Shibata,K., Shibata,Y., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (10-JUN-2000) Yoshihide Hayashizaki, The Institute of Physiological and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL: <http://genome-gsc.riken.go.jp/>, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT Please visit our web site (<http://genome-gsc.riken.go.jp/>) for

further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGGAGAAGCAGCGCCGCACATTCGCACTTTTGTATTATTAACC(CCCC(3') cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence[5' GAGGAGAAGCAGCGCCGCACATTCGCACTTTGTATTATTAACC(CCCC(3'] cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOLR.

FEATURES

source

Location/Qualifiers

1..2694

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="MGD:MGI:1907255"

/db_xref="taxon:10090"

/clone="1200017A24"

/sex="male"

/tissue_type="lung"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

22..1563

/note="Data source:SPT, source key:O9H103, evidence:ISS homolog to BA165P24.1.1 (NOVEL PROTEIN SIMILAR TO DROSOPHILA CG6630 AND CG11376, KIAA1058, RAT TRG (ISOFORM 1)) (FRAGMENT)"

putative"

/codon_start=1

/protein_id="BAB23587.1"

/db_xref="GI:12836283"

/translation="MOTPPPMQVEELCLNLSLYDYKMFQEDDEBMLDMLTRAKSYASADLR/LTWLQNNAEKHTKKCFETAAALCVAAALVAEITSLMEHSYLPVGGASTFONISNVLESAVSDDTLSPDEDVCSGRYFPESGLVLIDQLAEISTGELYEVNEYVKIVIPILFAHRDFRLKTSTDHKLOKAPDNINKDKRMGFYRGVGSRRYRGLDSEDEREYEKEPATTKLPETSHEEGFCGCGAFEEVIKDSTPVOKTKLDPRKAVAIQTPEPYEPDEYEEKDPVVYEFKNFLIRFMFTPTPLEGRPGIEHQHNRNVLLTMHAFFPKTRIRVRSOKREEVLTPIEAIEDMKKTKLQLAVATHOEPRPAKLQWLVLGSGVTANVGPLEVAOVPLAETPADPKYLHNHNLICRFERIMRGGAVERNRLTIASOREOUELKNKNYNLRMSLRPMIERKIPELYKPFRVDOSKRDSFISSRKCEYG LSOQS"

polyA_signal

2677..2682

/note="putative"

polyA_site

2694

/note="putative"

BASE COUNT

752 a 646 c 682 g 614 t

ORIGIN

alignment_scores:

Quality: 1821.50 Length: 471

Ratio: 4.296 Gaps: 2

Percent Similarity: 90.021 Percent Identity: 73.461

alignment_block:

US-09-737-246-2 x AK004816 ..

Align seg 1/1 to: AK004816 from: 1 to: 2694

1617 TTTAAGTAGUGUAAPRLeugLUteLAArgGIdUThrThePhProAspGlVa 1633
::: ||||| |
1 TTTCAGAGAGAGCACGCCCATCGACAACAACCTTTCCCATGCAGACT 50
1633 TGlnAspleuValPheasnleuHisMetileuleSerAsprthrValysm 1650
||||: |||:
51 GGAGAACTCTCTGCATCTGAACACACTTGTGTACGACACAGTGAGA 100
1650 eLyusGluHisGInGUasPrroGlmMeleuLeuIsprleuMytylmg 1666
||:::|| |
101 TAGGGAAATTCCAGGAAdAcCcTAgAttgGtttYmgGAGcStCaTgTrpCAGA 150

```

1667 1LeAlaLysGlyTyrGlnThrSerProGlu... ArgLeuThrTriPLeuL 1682
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 ATTGCCAAGAGCTACCGAGCATGCGCTGACCTGCGGCTGACTTGGCTCCA 200
1682 nasMetAlaGlyLysHisSerGluArgSerAsnHisAlaGluAlaAlaG 1699
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 GAACATGGCAGAGAACACACTAAGAAAGTGTCTCACAGAGGCCGCCA 250
1699 IncysLeuValHisSerAlaAlaLeuValAlaGluTyrLeuSerMetLeu 1715
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
251 TGTCTGTGGTGATGACGCCCTGTGGTGGCCAGTACCTGAGCATGTCTG 300
1716 GluAspArgLysTyrLeuProValGlyCysValThrPheGlnAsnIleSe 1732
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
301 GAGGACACAGACTGCTGCGGAGGCGGACGCTGACGTTTCAAAATTTTC 350
1732 rSerAsnValLeuGluGluSerAlaValSerAspValValSerProA 1749
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
351 TTCATATGTGCTTGAGAGCTGACACTCTGTATGACACTGTCTGACCTG 400
1749 spGluGluGlyIleCysSerGlyLysTyrPheThrGluSerGlyLeuVal 1765
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
401 ATGAGGACGCGGTATGCTGTGGTGTGCTGACTCTGATCTGAGAGTGGCTG 450
1766 GlyLeuLeuGluGlnAlaAlaAspPheSerMetAlaGlyMetTyrG1 1782
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
451 GCGCTCTGGAGAGCGCTGCGGAGCCTTCACAGCGGAGGAGCTGTACGA 500
1782 uAlaValAsnGluValTyrLysValLeuIleProIleHisGluAlaAsnA 1799
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
501 GACGGTTAATGAAGTCTACACAGCTGTCATCCCTATCTCGAGGACACA 550
1799 rAspAlaLysLysLeuSerThrIleHisGlyLysLeuGlnIleuAlaPhe 1815
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
551 GAAATTTCCGGAAGCTGACCTCCACTCAGCAGACAGCTCGAGAAGGCTTC 600
1816 SerLysIleValHisGlnSerThrGlyTyrGluArgMetPheGlyThrTy 1832
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
601 GATAACATCATCAACAAGAC.... CATAGAGATGTTTGGGAGCTA 644
1832 rPheArgValGlyPheTyrGlyThrLysPheGlyAspLeuAspGluGln 1849
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
645 CTTCACAGTGGTTTCTACGATCCCGATTTGGGATTTGGATGACAGAG 694
1849 LuPheValTyrLysGluProAlaIleThrLysLeuAlaGluIleSerHis 1865
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
695 AGTTCTGTACAAAGACCCGCAATCAGCAAGCTCCCGAGATCTCACAT 744
1866 ArgLeuGluGlyPheTyrGlyGluArgPheGlyGluAspValValGluVa 1882
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
745 AGACTAGAGGATTTATGGCCAGTGTTCGGTGACAGAGTTGTGGAAGT 794
1882 IleLysAspSerAsnProValAspLysCysLysLeuAspProAsnLysA 1899
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
795 GATAAAGACTCTACTCCAGTGGACAAAACCAAGTTGGATCTTACACAAG 844
1899 IaTyrIleGlnIleThrTyrValGluProTyrPheAspThrTyrGluMet 1915
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
845 CTTACTTTCAGATCACTTTGTGGAGCCTTACTTTGATGAATATGACATG 894
1916 LysAspArgIleThrTyrPheAspLysAsnTyrAsnLeuArgArgPheMe 1932
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
895 AAAGACCGGGTGACTACTTCGAGAAGAAATTTCAACCTCCGAGAGTTTAT 944
1932 TyrCysThrProPheThrLeuAspGlyArgAlaHisGlyLeuLeuHisG 1949
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
945 GTACACCAACCCCTTACCCCTGGAGGAGAGACCCGGGGCGAGCTTCATG 994
1949 LuGluPheArgLysThrIleLeuThrThrSerHisAlaPheProTyr 1965
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
995 AGCAACACCGCGAGAAACCGTGTCCACACCACTGACGCTTCCCTTAC 1044
1966 IleLysThrArgValAsnValThrHisLysGluGluIleIleLeuThrPr 1982

```

```

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1045 ATCAAGACAGGATCCGAGTACGAGAAAGAGGATTCGTTTGACTCC 1094
1982 oIleGluValAlaIleGluAspMetGlnLysLysThrGlnGluLeuAlaP 1999
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1095 GATTGAGGTTGCCAFTGAGATATGAGAAGAAACACCTCGACATGACCG 1144
1999 heaIaThrHisGlnAspProAlaAspProLysMetLeuGlnMetValLeu 2015
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1145 TGGCCACTCACAGAGAGCCCTGATGCAAAAGATGCTGCAAAATGTACTG 1194
2016 GlnGlySerValGlyThrThrValAsnGlnGlyProLeuGluValAlaG1 2032
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1195 CAGGCTCTGTAGAGACCACTGTAATCAGGACCACTGGAGGTGGCCCA 1244
2032 nValPheLeuSerGluIleProSerAspProLysLeuPheArgHisHisA 2049
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1245 AGTGTCTTGGCTGAATTCACAGTGAACCAAGCTCTACCGACATCACA 1294
2049 snLysLeuArgLeuCysPheLysAspPheThrLysArgCysGluAspAla 2065
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1295 ACAAGCTGAGGTTGTGCTTCAGAGAGTTCATTAATGCGATCGGAGAGCC 1344
2066 LeuArgLysAsnLysSerLeuIleGlyProValGlnLysGluTyrGlnAr 2082
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1345 GTGAGAGAAACAGAGCGACTCATACCCAGAGCAGCGGAGTACAGCA 1394
2082 gGluLeuGlyLys 2086
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1395 GAGCTGAGAAAG 1407
seq_name: gb_hc:BC020473
seq_documentation_block:
LOCUS BC020473 1510 bp mRNA linear HTC 03-JAN-2002
DEFINITION Homo sapiens, clone IMAGE:3853958, mRNA.
ACCESSION BC020473
VERSION BC020473.1 GI:18042973
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1510)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: villalob@bcm.tmc.edu,
Villalob, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
A.M., Holloway, W., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gunaratne, P., Yoon, V., Kowis, C., Martin, R.,
Lawrence, S., Richards, S., Gibbs, R.A.
FEATURES
source
1..1510
/organism="Homo sapiens"
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRK Plate: 13 Row: b Column: 12
This clone has the following problem: no polyA-tail.
Location/Qualifiers
1..1510
/organism="Homo sapiens"

```

```

/dB_xref="taxon:9606"
/clone="IMAGE:3853958"
/tissue_type="Colon, adenocarcinoma"
/clone_id="NH_MCC_65"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"

BASE COUNT      287 a      431 c      456 g      336 t
ORIGIN

alignment_scores:
    Quality: 1659.00      Length: 381
    Ratio: 4.621      Gaps: 0
Percent Similarity: 94.226      Percent Identity: 81.102

alignment_block:
US-09-737-246-2 x BC020473/rev ..

Align seg 1/1 to reverse of: BC020473 from: 1 to: 1510

1704 SerAlaAlaLeuValAlaGluTyrLeuSerMetLeuGlnAspArgLysTyr 1720
      ::::::::::::::::::::
1510 GGGGGCGCCCTGCTGGCTGAGTACCTCGCCCTGCTGAGAGACCAACCCTCA 1461
      ::::::::::::::::::::
1720 rLeuProValGlyCysValThrPheGlnAsnLeuSerSerAsnValLeug 1737
      ::::::::::::::::::::
1460 CCTGCCCGTGGGTGGTTCCTTCCAGAACATCTCATCATGTGCTAG 1411
      ::::::::::::::::::::
1737 InGluSerAlaValSerAspAspValValSerProAspGluGluGlyIle 1753
      ::::::::::::::::::::
1410 AGGAGTCCGCCAATCTCCAGAGACATCTGTCGCCCGCCGAGGAGGGCTTC 1361
      ::::::::::::::::::::
1754 CysSerGlyLysThrPheThrGluSerGlyLeuValGlyLeuGluGluGlu 1770
      ::::::::::::::::::::
1360 TGTCTCGGGAGACACTTCACTGAGCTGGGGCTGTGAGGGTTGCTGGAGACA 1311
      ::::::::::::::::::::
1770 nAlaAlaAlaSerPheSerMetAlaGlyMetTyrGluAlaValAsnGluV 1787
      ::::::::::::::::::::
1310 GGCAGCGCGCTACTTCAACATGGGCGGGCTCTACGAGCGGTGAAATGAGG 1261
      ::::::::::::::::::::
1787 aLlTyrIlyValLeuIleProIleHisGluAlaAsnArgAspAlaLysLys 1803
      ::::::::::::::::::::
1260 TCTACAGAAGAACCTCAATCCCATCTCTGGAAACCCACCGTGACTACAAAG 1211
      ::::::::::::::::::::
1804 LeuSerThrIleHisGlyLysLeuGluGlnAlaPheSerLysIleValAl 1820
      ::::::::::::::::::::
1210 CTGGCCGCGGTGACGCGCAACACTGCGAGAGGCTTCAACCAAGATCATGCA 1161
      ::::::::::::::::::::
1820 sGlnSerThrGlyTyrPgluArgMetPheGlyThrTyrPheArgValGlyP 1837
      ::::::::::::::::::::
1160 CCAGAGTTCGCGGTGGAGCGCGCTGTTCGGAGAGTATTTCCGGTGGGCT 1111
      ::::::::::::::::::::
1837 heTyrGlyLthrTyrPheGlyAspLeuAspGluGluGlnGluPheValTyrLys 1853
      ::::::::::::::::::::
1110 TCTACGCGCCCACTTCCTGCTTGCCTGATAGAGAGGTTGTGTACAG 1061
      ::::::::::::::::::::
1854 GluProAlaIleThrTyrLeuAlaGluIleSerHisArgLeuGluGlyPh 1870
      ::::::::::::::::::::
1060 GAGCCATTCATCCATCGAAGCTGCGAGAGATTCACACCGGCTGGAGAGCTT 1011
      ::::::::::::::::::::
1870 eTyrGlyGluArgPheGlyGluAspValValGluValIleLysAspSerA 1887
      ::::::::::::::::::::
1010 CTACACGAGAGAGATTGGCGACGACGTCGTGAGATTATCAAGAAGCTCA 961
      ::::::::::::::::::::
1887 snProValAspLysCysLysLeuAspProAsnLysAlaTyrIleGlnIle 1903
      ::::::::::::::::::::
960 ACCCTGTGACAAGTCCAAAGCTTGACTCACAAAAGGCTCATCTCAAGATC 911
      ::::::::::::::::::::
1904 ThrTyrValGluProTyrPheAspThrTyrGluMetLysAspArgIleTh 1920
      ::::::::::::::::::::
910 ACCTATGTGGAACCGTACTTTGTACTACGAGCTCAAGAGCCGGGTGAC 861
      ::::::::::::::::::::
1920 rTyrPheAspLysAsnTyrAsnLeuArgArgPheMetTyrCysThrProP 1937

```

|||||..... |||||
860 CTACTTTGACCGCACTATGGCTTCGCACATTCCTGTTCGCACCCT 811
1937 hehrlieuhsprgluargalnahisglyluenhisgluinpheylsavg 1953
810 TCACGCCGATGGCGCACAGGGGACCTCCGCACAACAAGCST 761
1954 lvsThrlleuthrthrserHisAlaphrotyrFllysthrarvya 1970
760 AAGACCTCTGCAGCACGACACGCGCTTCCCTACATCAAGACTCGCAT 711
1970 lnsnvalThrHslsgluGlulIelleuthrProileguValAlal 1987
710 CCGTGCTGCCACCGGGAGAGAGCGTCTGACAGCCAGTAGAGTGCGCA 661
1987 legluaspmetcInlysThrngluLeualaAhealathrhlsin 2003
660 TCGAGGACATGCGGAAGAAGACAGGGGACCTGCGCTTCCACCGGAG 611
2004 AspPrclAspprolysmeltleugImetValleugllySerValGI 2020
610 GACCCACGAGATCTTAAGTCTACAGATGTGCTTCAGGGCTCTAGG 561
2020 yThrThrValasnglnglyProleuguValAlaglnValPheLuserg 2037
560 GCCCACCGGAAACAGAGTCCCCTCGAGGTGGCCCCAGTGTTTTAGCAG 511
2037 InuleproserAspprolysleupharArgHisInslnysleuArgleu 2053
510 AGATCCCGGAAGACCCCAAGCTTCCCGCATCACAAATTGGCGGCTC 461
2054 CyAPhelaAspPherThrlysrArgCysgluAsPALeuarqlyslny 2070
460 TGCTTCAGAGACTTCGCAAGAATGTGAGAGATGCCCTCGCGAAAATA 411
2070 sSerleuIlleglyProvalgluLysglutrglnArvgIleu 2084
410 GGCCCTGATTGGCGCGAGACAGAAAGATACCCGCTGAGCTG 368
seq_name: gb_htc:AKO10755

seq_documentation_block:
LOCUS AKO10755 1507 bp mRNA linear HTC 19-JAN-2002
DEFINITION Mus musculus ES cells cDNA, RIKEN full-length enriched library,
clone:2410095B20;homolog to KIAA1395 PROTEIN (FRAGMENT), full
insert sequence.
ACCESSION AKO10755
VERSION AKO10755.1 GI:12846421
KEYWORDS HTC; GAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) ES cells cDNA to mRNA,
clone.lib:RIKEN full-length enriched mouse cDNA library
clone:2410095B20.

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Carninci,P., Itoh,M., Hayatsu,N., Sugahara,Y., Shibata,K.,
1 (sites)
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
2 (sites)
REFERENCE Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
AUTHORS Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
3 (sites)
REFERENCE Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
AUTHORS Kono,H., Akiyama,J., Nishi,K., Katsuna,T., Tashiro,H., Itoh,M.,
Sunil,N., Ishii,Y., Nakamura,S., Hazama,M., Niehlne,T., Harada,A.,

Yamamoto, R., Matsunoto, H., Sakakuchi, S., Ikegami, T., Kasaiiwa, K., Fujiwara, S., Inoue, K., Togawa, Y., Irawa, M., Ohara, E., Matsubiki, M., Okazaki, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system - 384 format sequencing pipeline with 384 multiplexillary sequencer genome Res. 10 (11), 1757-1771 (2000)

TITLE	JOURNAL	PUBMED	REFERENCE
1. The effect of the use of a...	Journal of Clinical Medicine	123456789	1. Smith J. The effect of the use of a...
2. The effect of the use of a...	Journal of Clinical Medicine	123456789	2. Smith J. The effect of the use of a...
3. The effect of the use of a...	Journal of Clinical Medicine	123456789	3. Smith J. The effect of the use of a...
4. The effect of the use of a...	Journal of Clinical Medicine	123456789	4. Smith J. The effect of the use of a...
5. The effect of the use of a...	Journal of Clinical Medicine	123456789	5. Smith J. The effect of the use of a...
6. The effect of the use of a...	Journal of Clinical Medicine	123456789	6. Smith J. The effect of the use of a...
7. The effect of the use of a...	Journal of Clinical Medicine	123456789	7. Smith J. The effect of the use of a...
8. The effect of the use of a...	Journal of Clinical Medicine	123456789	8. Smith J. The effect of the use of a...
9. The effect of the use of a...	Journal of Clinical Medicine	123456789	9. Smith J. The effect of the use of a...
10. The effect of the use of a...	Journal of Clinical Medicine	123456789	10. Smith J. The effect of the use of a...

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE
JOURNAL
REFERENCE

5 (Pages 10 to 1507)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
Hara, A., Hayata, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,
Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koyas, S.,
Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
Numata, K., Ono, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Schraml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tatemai, M., Tagawa, A., Takahashi, F.,
Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I.,
Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and
Hayashizaki, Y.

TITLE
JOURNAL

Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: riken-ies@gsc.riken.go.jp,
<http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5', GAGGAGAGAGAGATCAAGACGCTCTTTTTTTTTTTTAA 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went by through one round of normalization to Rot - 5.0 and subtraction to Rot - 25.0. Second strand cDNA was prepared with the primer adapted of sequence [5', GAGGAGAGATCTCGACTTAATTAAATTAATCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. host: SOLR.

FEATURES

```

/cell_type="ES cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
600. .1433
CDS

```

```

/codon_start=1
/protein_id="BAB27161.1"
/translation="MEDEEPLASALRIASPEVIAVSHLVIVLTVLPVSPITCGG
PVTQVATLRGSRGSRVLTAKSKITSNDPLAVPSSVDVEVRLTAAHAFQDLMV
MYNIGRAFPAMAHVASLVHRLNEAVQDSRGCHPLASTVYHAFPLPGDSLREPAH
PRAVSYTEASSSAVPAQSHRYQKLHEELALQWVSSASRELVLQAHAFQDLMV

```

BASE COUNT	269 a	490 c	413 g	335 t
ORIGIN	KSMLEHLILGORDTPRNVRPGRRFLDDIALVASGLEVITRHKTKGCLCCTLGCCLA			

alignment_scores

Quality:	1396.50	Length:	500
Ratio:	3.474	Gaps:	4
Percent Similarity:	80.400	Percent Identity:	53.200

```
alignment_block:
```

Align seg 1/1 to: AK010755 from: 1 to: 1507

567 YRPRPGIINSEIRLEASNPhEalAsnAhrgrLIInLysERLAaTgrsLI 583
|||||.....
3 TACCAACACAGCCTCAATTTACAGCAAGTCGCGAGGGCTCTGTGGCCAACCT 52
583 eRhYAlLysVaLIgInPhEmETyrgLIgLIaSProSErAsnAlaMeP 600
.....
53 GGCTGGGGGATCCAGTACATGGCGGGGAAGACACAGCGCAGCCCTTGC 102
600 roVAlIIEPhEgLIyLysSErSercySgrLIuPhEsrLIgLIaLaTyR 616
|||||.....
103 CGGTATCTTTGGGAATCTAGCTGCACAGCAATTCACCCGAGAGCCCTTC 152
617 ThralaVaLIaTyHIsAsnArgSErProaPhEnIsHIsLIeulE 633
|||||.....
153 ACACAGAGTGCTATACACACAAAGTCTCTGAATTTACAGAGCAATTCAA 202
633 sValLysLEuProaLarhLEuThrAsPnIsHIsLIeulEuphEThr 650
I.....
203 GCTACAGACTCTCGCTCGCTCGTACCGAGAACCACTACCTCTTTTACCT 252
650 hEThYHIsVaLSErCySgInLIgLIaSPnThrProLEuGInThrPro 666
|||||.....
253 TGTACCAATGTACAGTCCAGCCGCGCCGAGCAACAGCTTTGAGACCCCT 302
667 ValgLyTyThrTrPIeProMEtLEuGInsnLIaTyrgLEuLysThrgI 683
|||||.....
303 GGGGGCTTCACTGGATCCCTGTGTACAAATGGCGCGCTGAGGACTGG 352
683 yGInPhEcYsLEuProVaLSErLEuGILySProGInLaLaTySErV 700
I.....
353 TCCCTTGCCTGGCCCGTCCGTCCGTGGACACACCTCCACCCAGCTACCTGG 402
700 alLEuSErProGInVaLProLEuProGIyMEtLSTrVaLAsPnHIs 716
|||||.....
403 TCTGTACCCCAATGTACGCTGCCCTGGCATGCCCTGGGTGGAGGCCAC 452
717 LysGIyVaLPhEasnVaLIgLIaVaLaLaVaLSErSeriLEuThrgI 733
|||||.....
453 AAGGCGGTGTACGTGTGAGACTACAGCCGCTGTGCCTCGACACCCGCA 502
733 nAsPProTyLEuAsPlySPhePhEaLaLEuVaLAsnAlaLEuAsPlyuH 750
|||||.....
503 GGACCCCACTTGATGAATGTTCTTACGCGTGTCATGTTCTGTAGAGGAAG 552
750 IsLEuPhEProVaLarGLIEgLIaSPheTyrgLIEMeTGIuAsnAsLEu 766
.....
553 GGATCTTTCCTCGTCCGACTCAAGCAGACAGTGCAGCTTAAGGCACATAG 602
767 GIuAsnGILeULySErSErTIESErAlaLEuAsnSErSErGInLEuGI 783
|||||.....
603 GAGCAGGAGTTGGCGAGCCAGCCTGGAGCCCTGGCCTTCCAGACCCAGA 652
783 uProVaLArGPhELeuHIsLEuLEuLEuLEuAsPlyLEuLEuLEu 800
.....
653 GCCCTAGTACGCTTTTCCCACTCGTGTAGACAAACCTGTGCGCTGG 702
800 alLEaRgProRoVaLIIEaLagLIyInLIeVaLsnLEuGIyGInLa 816

```

|||||
703 TTGTGGGGCCACCATCATTTGTGGCCAGATGTGAACCTGGGTGAGGG 752
817 SerPheGluAlaMetAlaSerIleLeuHisArgLeuHisLysAsnLeuGlu 833
|||||
753 GCCTTTGAAGCTATGCTCATGTATGACCAACCTTGTGCACCGCAACCTGGA 802
833 uGlyAsnHisAspGlnHisGlyArgAsnSerLeuLeuAlaSerTyrIleH 850
|||||
803 GGCTGTCAAGACTCCCGTGGCCACTGGCCACTGCTGGCTCTATGTCC 852
850 iSTyrAlaPheArgLeuProAsn.....ThrTyrProAsnSerSer 863
|||||
853 ACTAGCCCTTTCGCCCTGTGTGTGACCTTCAGCTGCACGAGCGAAGCC 902
864 SerProGlyProGlyGlyLeuGlySerValHisTyrAlaThrMetAl 880
|||||
903 CCGCCA.....GCAACTGTGCAGGCGCCACACTGGC 934
880 aArgSerAlaValArgProAlaSerLeuAsnLeuAsnArgSerArgSerL 897
|||||
935 CCGTGGCTGTGTGGCCCGCCAGCCTGTACTGTGCACGCTTAAGAGCA 984
897 euSerAsnSerAsnProAspIleSerGlyThrProThrSerProAsp 913
|||||
985 TCAGCAGCAGCAACCTGACTGGCTGTGGCTGTGGCTGTGGCAAT 1034
914 GluValArgSerIleIleGlySerLysGlyLeuAspArgSerAsnSerTr 930
|||||
1035 GAGCTGTCCCGCATCTCTAGCAGCAAGGCAATGCACCGCAACTGATAC.. 1082
930 pValAsnThrGlyGlyProLysAlaAlaProThrGlySerAsnProSerP 947
1082 ..... 1082
947 roSerAlaGluSerThrGlnAlaMetAspArgSerCysAsnArgMetSer 963
1083 .....CGGGCTCC 1091
964 SerHisThrGlu...ThrSerSerPheLeuGlnThrLeuThrGlyArgLe 979
|||||
1092 TCTTACTCGAGGCTCATACCTCGCGCCGCCAGCCCTCAGTCAGAGACA 1141
979 uProThrLysLysLeuPheHisGlnGluLeuAlaLeuGlnThrPheVala 996
|||||
1142 CACTGTGCAGAAAGCTCTGTCAGCAGAGAGCTGGCTGTGCAGTGGTGTGA 1191
996 ysSerGlySerValArgGluSerAlaLeuGlnGlnAlaThrPhePhePhe 1012
|||||
1192 GCGGCAAGCGGCTTCGTGAGCTAGTCTCTGACAGATGCTGTCTTCTTC 1241
1242 CAGCTCATGTGTAATAAGCATGAGCTTCATCTGCTCTGGGCGACGAGACT 1291
1029 uGluAlaProArgLysSerArgPheProGluArgPheMetAspAspIleA 1046
|||||
1292 GGACACTCCCGCAGACGTGGCTTCCCTGGCGCTTCTGTGAGAACATTGG 1341
1046 laaIalaValSerThrIleAlaSerAspIleValSerArgPheGlnLys 1062
|||||
1342 CCGGCCCTGGGTGCTCCGTGGGCTGGAAGTCAACACAGATCCATTAAG 1391
seq_name: gb_est1:AJ397707
seq_documentation_b1007:
LOCUS AJ397707 796 bp mRNA linear EST 25-JAN-2001
DEFINITION AJ397707 dkfz426 Gallus gallus cdna clone 30k15r1, mRNA sequence.
ACCESSION AJ397707.1 GI:7130918
VERSION AJ397707.1
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauromia; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 796)
AUTHORS
Abdrakhmanov, I., Lodgyn, D., Geroch, P., Arakawa, H., Law, A., Plachy
, J., Korn, B. and Buerstedde, J.M.
A large database of chicken burasal ESTs as a resource for the
analysis of vertebrate gene function
Genome Res. 10 (12), 2062-2069 (2000)
JOURNAL
MEDLINE
20568495
COMMENT
Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinstr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
Location/Qualifiers
FEATURES
source
/organism="Gallus gallus"
/strain="CG"
/db_xref="taxon:9031"
/clone="30K15r1"
/clone_1ib="dkfz426"
/issue_type="Bursa of Fabricius"
BASE COUNT 210 a 186 c 202 g 198 t
ORIGIN
alignment_scores:
Quality: 1293.50 Length: 265
Ratio: 4.937 Gaps: 1
Percent Similarity: 98.868 Percent Identity: 96.981
alignment_block:
US-09-737-246-2 x AJ397707 ..
Align seg 1/1 to: AJ397707 from: 1 to: 796
1482 ValGlnThrValSerValThrGluSerLysGluSerIleLeuGlyValA 1498
|||||
1 GTACGAGCTGTTTCTGTGACGAGTCCAAAGAGATATCCTGGCGGGGT 50
1498 lleuLysValleuLeuHisSerMetAlaCysAsnGlnSerAlaValTyrL 1515
|||||
51 GCTGAGGCTGCTTCTGTCACACAGATGCGCTGCACCAAAAGTCACTTATC 100
1515 euGlnHisCysPheAlaThrGlnArgAlaLeuValSerLysPheProGlu 1531
|||||
101 TCCACACTGCTTCCGCACACAGAGAGCTTGTCTCAAAAGTTCCCTGAG 150
1532 leuLeuPheGlnGluGluThrGlnGlnGlnCysAlaAspLeuCysLeuArgLe 1548
|||||
151 CTGCTGTTGAGAGAGAGAGAGTGCAGTGCAGATCTGTCTCTGAGGCT 200
1548 uLeuArgHisCysSerSerSerIleGlyThrIleArgSerHisProSerA 1565
|||||
201 CCGAGGACATCTATACACAGATGACACAAATAGCGTCAACACCGAGATG 250
1565 laSerLeuTyrleuLeuMetArgGlnAsnPheGlnIleGlyAsnAsnPhe 1581
|||||
251 CCTCACTTACCTACTCATGTAGGCAAAACTTGTGAGATTGGCAATTAACCTC 300
1582 AlaArgValLysMetGlnValProMetSerLeuSerSerLeuValGlyTh 1598
|||||
301 GCCAAGATGAATAATGACAGTACTATGTCTCTCTCTGCTCCGTGGGGAC 350
1598 rSerGlnAsnPheAsnGlnGluPheLeuArgArgSerLeuLysThrIleL 1615
|||||
351 ATCCACGAAATTTCAATGAGGAATTTCTACAGCTTCCTTGAAGATATTT 400
1615 euThrTyrAlaGluLysPheGlnGluLeuArgGlyThrThrPheProAsp 1631
|||||
401 TAACGATGCTGAAGAGATCTGGAACCTTAGAGACAAACGTTTCTGAT 450

```

1632 GlnValGlnAspLeuValIlePheAsnLeuHisMetIleLeuSerAspTrpVal 1648
|||||
451 CAGGTCCAAAGATCTGGTTCACCTCCACATGATCTCTCTGACACAGT 500
1648 llysMetLysGluHisGlnGluAspProGluMetIleAspLeuMet 1665
|||||
501 TAAATGAAGACGACACGAGAGATCCAGAAATGCTATATTGACCTGATGT 550
1665 yratGilealalyGlyTyrGlnThrSerProGlu...ArgLeuThrTrp 1680
|||||
551 ACAGGATGGCAAAAGGCGCTATCAGAAATTCCTCCCTACCGCCCTTGACCTGG 600
1681 LeuGlnAsnMetAlaGlyLysHisSerGluArgSerAsnHisAlaGlu 1697
|||||
601 TTTCAGACACATGGCTGCAAAAGCATCGGAGACAGACATATGACAGATC 650
1697 aalagGlnCysLeuValHisSerAlaAlaLeuValAlaGluTyrLeuSerM 1714
:|||||
651 AGCAGATGTTTGTGTCACCTCGACGCTTGCTGGCTGAGTATCTGACGA 700
1714 etleuGluAspArgLysTyrLeuProValGlyLysValThrPheGlnAsn 1730
|||||
701 TGCTCGAGGACCGCAAAATACCTCCCTGTAGAGCGCTTACATTTTCAGAAC 750
1731 lIleSerSerAsnValIleuGluGlnSerAlaAlaValSerAspAspVal 1745
|||||
751 ATATCTTCGAATGTTTGGAAAGTCAGACGATTTTGATGATGATGTC 795
seq_name: gp_hnc:AK018051
seq_documentation_block:
LOCUS AK018051 1712 bp mRNA linear HTC 19-JAN-2002
DEFINITION Mus musculus adult male thymus cDNA, RIKEN full-length enriched
library, clone:5830472H07:homolog to KIA1395 PROTEIN (FRAGMENT),
full insert sequence.
ACCESSION AK018051
VERSION AK018051.1 GI:12857594
KEYWORDS HTc; CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male thymus cDNA to mRNA,
clone:lib:RIKEN full-length enriched mouse cDNA library
clone:5830472H07.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
JOURNAL 99279253
MEDLINE 10349636
PUBMED
2 (sites)
REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Komuro,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
to Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL 20499374
MEDLINE 11042159
PUBMED
3 (sites)
REFERENCE
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Komuro,H., Akiyama,D., Nishi,K., Kitunai,T., Tashiro,H., Itoh,M.,
Suni,N., Ishii,Y., Nakamura,S., Hazama,M., Nishise,T., Harada,A.,
Yamamoto,R., Matsunoto,H., Sakaguchi,S., Ikegami,T., Kasaiwa,Y.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M.,
Toneeda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunaga,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
to Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL 20530913
MEDLINE 11076661
PUBMED
4 (sites)
REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the

[illegible]

57 CTTTATGAGCCTGTGAGCCAGCTTGAAGACTGTGATACACA 106
67 sproleuAlaValaspserGlyProleuAargspLeuIleGluPhePro 84
107 CTTAAACACCTTGATGACAGCTGCCCAAGAGCTGGGGACCTCACCG 156
84 roAspAspIleGluValTyrserProAArgspCysArgThrIleVal 100
157 ATGACGACCTGCATGTGGCTTACACACCAAGATGTAGACTTTGCAG 206
101 SerAlaValProGluGluSer...GluMetAspProHisValArgAspCys 116
207 CACTCTCTCCAGAGAGAGAGTGAATCGATCTCAGCTCAGAGACTG 256
116 sIleAArgserTyThrGluAspThrAlaIleValIleArgIleTyThrIst 133
257 TGTTCAGACCTTATTTGAGAGTGGCTGATTTGAACCGGAAAAACAG 306
133 ySleuGlyThrGlyPheAsnProAsnThrLeuAspIlysgIlnysGluArg 149
307 AAGTTCAGAGTTTTCAGCTTTAAAGACGGGATCTCGACAGATTTT 356
150 GluIlysgIlyLeuProIlysgIlnValPheGluSerAspGluAlaProAspCyl 166
357 CAGAGACCTTTCAGAAACAGACGTTTGAGTCAGAAACCTTGAGTGCAG 406
166 y.AsnserTyrgIlnAspAspGlnAspAspLeuIlysgArgSerMetSer 182
407 TGAAC.....CGACACTCAGC 423
183 IleAspAspThrProArgIlySerThrAlaCysSerIlePheAspLeu 199
424 CTCGAG..... 429
199 saAsnSerLeuProAspAlaLeuLeuProAsnLeuLeuAspArgThrPro 216
430CCTGATGAGCGGCTGGAACCTGCTCCAGCTTGAGCG 469
216 snGluGluIleAspArgGlnAspAspAspGlnArgIlySerAsnArgHis 232
470 CTGAGGACTTGTGAAGAGAGAGAGAGGCCCGCAAGACCAATTCGCGC 519
233 LysGluLeuPheAlaLeuHisProserProAspGluGluGluProIleG 249
520 GGTGAGCTTTTGCCTATCCGCCCGATGAGAGAGATCTGTGGA 569
249 uArgLeuSerValProAspIleProIlysgIlnHisPheGlyGlnArgLeu 266
570 AATACGTCAGTACATGAATGCTTAAGCAACATCTGGGCAACAGAAATAT 619
266 euValIlysgIlySerLeuIlysgIlnPheGluIleGluIleGluProIlePhe 282
620 TGGTCAAGGTGCTGACTGTAAGTTTGAGATTGAATGAACCTGTGTT 669
283 AlaSerLeuAlaLeuTyrsValIlysgIlnLysIlysgIlnSerGluAs 299
670 GCGAGTGTTCCTCTATGACGTTAAAGAAAGAAAAAGATCTCAGAAAA 719
299 nPheTyrsPheAspLeuAsnSerGluGluMetIlysgIlyLeuLeuArgPro 316
720 TTTCCACTGTGACTGAATCCAGACAGTCAAAAGGTTTCTCGAGCTC 769
316 IsValIleProAlaAlaIleThrThrLeuAlaArgSerAlaIlePheSer 332
770 ACACACCTCGATTGACCCATCGAGTGAAGCTGAGTGTGCGGTCTCT 819
333 IleThrTyrsProserGlnAspValPheLeuValIleIlysgIlnLysVal 349
820 GTCCACCAACCATCTTCAGACATCTACTTGGTGTCAAGATGAAAAAGT 869
349 IleuGlnGlnGlyAspIleGlyIlyGlyGlyGlyGlyGlyGlyGlyGly 366
870 CCTTCAGCAAGAGAGATGTGACAGTGTGCAAGAACCTTACATGATCATCA 919

366 ySgIlnAlaAspAlaThrLysAsnIlysgIlnLysLeuIlysgIlnSer 382
920 AAGAAGACGATGTGTAAGAGTAAAGAAAGTGAAGAAAAAACAATTA 969
383 GluAlaAspGlnPheCysGlnArgLeuGlyIlyTyrsArgMetProPheAl 399
970 CAACGTGATCTTCTGCAACGTTTGGGAAATATACCGGATGCTTCCG 1019
399 aTyrsThrAlaIleHisLeuMetAsnIleValSerSerAlaGlySerLeu 416
1020 CTGGGCCCCCTTGTAGCTTACCAACCTTCTTC...AACATCTCCACCTTG 1066
416 LuArgAspSerThrGluValGluIleSerThrIlysgIlnArgIlysgIlySer 432
1067 AAAGGAGACGACAGATGTGAGCGCTGGGTTGGC...AGAACTCT 1110
433 TrpSerGluArgArgAsnSerSerIleValGlyArgArgSerLeuGluArg 449
1111 GTGGGTGAGAGAGAGAGCTGTCTCC...CAATCCAGAGAGGCCCTCTGAGCG 1157
449 gThrThrSerGlyAspAspAlaCysAsnLeuThrSerPheArgProAla 466
1158 AACCTCTCTCTTGGAGAAATGAGATTGATCCGACTTCAAAACCCACA 1207
466 hrLeuThrValThrAsnPhePheIlysgIlnGluIlysgIlnArgSerAsp 482
1208 CCTTGCCACCAACATCTTCTTCAACAGAGAGAGATCGCTTGAATGAT 1257
483 GluAspLeuTyrsPheLeuAlaAspMetArgArgProserSerValLe 499
1258 GAGACTGTGTCAAGTTTGTAGCTGACTACAAAGAGATCTTCATCCACCA 1307
499 uArgArgLeuArgProIleThrAlaGlnLeuLysIleAspIleSerPro 516
1308 GGAAGAGTCAATCATCCAGGCTCAGTGAAGTGAAGTATATCCACAG 1357
516 IaProGluAsnProHisTyrsCysLeuThrProGluLeuGlnValLys 532
1358 CTCGCCGAGCTGATGAACTGCTGCTGAGACGCCGAGGTCTCCAGTCAAA 1407
533 LeuTyrsProAspSerArgValArgProThrArgIleGluLeuGluPhe 549
1408 CCTTTCCTGAAATCGGAGCGGTCCACCAAGAGATTTTGAAATTTCC 1457
549 oAlaArgAspValTyrsValProAsnThrThrTyrsArg.....A 562
1458 GATCCGGAGAGGTGACGTCCCTCACACGTGTGACAGGTGAGAAATGTGGG 1507
562 snLeuLeuTyrsIleTyrsProGlnSerLeu 571
1508 CCTTGTGTGATATACATCTGCTCTC 1536
seq_name: gb_est2:B6247108
seq_documentation_block: 931 bp mRNA linear EST 13-FEB-2001
LOCUS B6247108
DEFINITION B02359776F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488235 5',
mRNA sequence.
ACCESSION B6247108
VERSION B6247108.1 GI:12756923
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 931)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Published (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: gcapso-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.

CDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: L1AM1034 row: 9 column: 20
 High quality sequence stop: 728.

FEATURES
 source
 Location/Qualifiers
 1..931

/organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:448235"
 /clone_lib="NCI CGAP Mam1"
 /tissue_type="tumor biopsy sample"
 /dev_stage="10 months, virgin"
 /lab_host="DHI0B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Salt;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"
 BASE COUNT 275 a 198 c 228 g 230 t
 ORIGIN

alignment_scores:

Quality: 1209.50 Length: 272
 Ratio: 4.670 Gaps: 4
 Percent Similarity: 95.221 Percent Identity: 88.971

alignment_block:

US-09-737-246-2 x BG247108 ..

Align seg 1/1 to: BG247108 from: 1 to: 931

1742 Seraspapvalvalserproaspplugluclylecysserglylsty 1758
 ||||||||||||||||||||||||||||||||||||||||||||
 2 TCAGATGATGTGTCATCCAGATGAGAGCATTTCTCTCGAAGTCA 51
 ||||||||||||||||||||||||||||||||||||||||||||
 1758 rphethrgluserglyleuvalgyleuleuglinalalaalaserp 1775
 ||||||||||||||||||||||||||||||||||||||||||||
 52 CTCACCTGATGGGAGCTGGGCTTCTGGAACAACGGCTGCTTCT 101
 ||||||||||||||||||||||||||||||||||||||||||||
 1775 hesermetalsglymetyrglualalaasngluvaltyrlyslaleu 1791
 ||||||||||||||||||||||||||||||||||||||||||||
 102 TCTCATGGCGGGCATGTACAGACGATTATGAGTTTACAAAGTATC 151
 ||||||||||||||||||||||||||||||||||||||||||||
 1792 illeprollehsglualalasnarqaspaalalsylsleuserthrllehi 1808
 ||||||||||||||||||||||||||||||||||||||||||||
 152 ATTCTATTTCATGAGCTAATCGGGATCCAAAGAGCTCTCCCAATTC 201
 ||||||||||||||||||||||||||||||||||||||||||||
 1808 sglylsleuglinalapheserlyllevalhisglinsertbrglyt 1825
 ||||||||||||||||||||||||||||||||||||||||||||
 202 TGGCAACTTCAGAGGCGATTCAGCAAAATGCTTCATAGGAT.....G 245
 ||||||||||||||||||||||||||||||||||||||||||||
 1825 rpgluarqmetepheglythrtyrphearqvalgllypheyrglythrllys 1841
 ||||||||||||||||||||||||||||||||||||||||||||
 246 GTAACCGGATGTTGGCACCTATTCCGTGTGGTTTATGAAACCAAG 295
 ||||||||||||||||||||||||||||||||||||||||||||
 1842 pheglyaspleuaspluglulphvaltyrlysgluoproalalietrh 1858
 ||||||||||||||||||||||||||||||||||||||||||||
 296 TTCGGGGATTGGATGAGACAGAAATTGTTTACAAAGAGCCTGCACATTAAC 345
 ||||||||||||||||||||||||||||||||||||||||||||
 1858 rlyslenuaglulieserhisargleuglulglypheyrglygluarqrp 1875
 ||||||||||||||||||||||||||||||||||||||||||||
 346 CAATCTCCAGAGATTTCTCACAGATTGGAGGATTTTATGAGAAAGAT 395
 ||||||||||||||||||||||||||||||||||||||||||||
 1875 heglygluaspyalvalgluvalilleyaspsersasnprovalasplys 1891
 ||||||||||||||||||||||||||||||||||||||||||||
 396 TTGGAGAGATGTGCTGAAGTAATCAAGGACCTAATCCGTGACAGAG 445
 ||||||||||||||||||||||||||||||||||||||||||||
 1892 Cyslysluaspproasnlyslatyrilleglillethrtyrvalgluipr 1908

||||||||||||||||||||||||||||||||||||||||||
 446 TCGAATTAGATCCAAACAGGCAATATATTCAGATTCTACGTGGAGCC 495
 ||||||||||||||||||||||||||||||||||||||||||||
 1908 otyr-pheaspthrtyrglumetlysasparqilethrtyrpheasplysa 1925
 ||||||||||||||||||||||||||||||||||||||||||||
 496 CTCTCTTGACACACTATGATGAGTGAAGACAGATCACTTATT. GACAAA 544
 ||||||||||||||||||||||||||||||||||||||||||||
 1925 snlyrasnleuarqarphemetlyrcysthrprophetrleuaspgly 1941
 ||||||||||||||||||||||||||||||||||||||||||||
 545 ATTATTAATCTCCGGCGTTTCATGTACTGCAACACCTTCACCTTAGATGCG 594
 ||||||||||||||||||||||||||||||||||||||||||||
 1942 Argalahisglulguhlsiglulphelysarqysrhrilleuthr 1958
 ||||||||||||||||||||||||||||||||||||||||||||
 595 CGTCTCATGGGAGCTTCATGACAAATTCACAAAGAGACCATTTCTGAC 644
 ||||||||||||||||||||||||||||||||||||||||||||
 1958 rthserhis.alapherprotyrilleysthrarqvalasvalthrhis 1974
 ||||||||||||||||||||||||||||||||||||||||||||
 645 AACCTCATGTCTCTTCTTACATTAACAAAGAGTCATGTCACGTAC 694
 ||||||||||||||||||||||||||||||||||||||||||||
 1975 lysglulguillethleuthrprollegluvalalalegluaspmetgl 1991
 ||||||||||||||||||||||||||||||||||||||||||||
 695 AAAGAGAGATTAATTAACCAATGGAAGGTCTATCCGAGACTGCA 744
 ||||||||||||||||||||||||||||||||||||||||||||
 1991 nlyslsThrnglulduvalaphaalathrhls.glnaspproalaasp 2007
 ||||||||||||||||||||||||||||||||||||||||||||
 745 GAAGACAGCCCGA.....ATGGCTTGACACACATATGTCGCCGGA 785
 ||||||||||||||||||||||||||||||||||||||||||||
 2008 Prolysmetleu 2011
 ||||||||||||||||||||||||||||||||||||||||||||
 786 CCAAGGTAATG 797
 ||||||||||||||||||||||||||||||||||||||||||||
 seq_name: gb_hlc:BC018075

seq_documentation_block:

LOCUS BC018075 1873 bp mRNA linear HTC 06-DEC-2001
 DEFINITION Homo sapiens, clone IMAGE:4792506, mRNA.
 ACCESSION BC018075
 VERSION BC018075.1 GI:17390162
 KEYWORDS HTC.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 1873)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
 Toshitsugu and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) md@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

COMMENT

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAC Plate: 32 Row: b Column: 21
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Similarity but not
 identity to protein
 This clone has the following problem: frame shifted.

FEATURES

source
 1..1873
 Location/Qualifiers

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4792507"
/tissue_type="Brain, hypothalamus"
/clone_id="NIH_MGC_96"
/lab_host="DH10B"
/notes="Vector: pBluescript"
BASE COUNT      604 a      317 c      365 g      587 t
ORIGIN

```

```

alignment_scores:
  Quality: 1199.00      Length: 229
  Ratio: 5.259          Gaps: 1
Percent Similarity: 99.563      Percent Identity: 99.127

```

```

alignment_block:
US-09-737-246-2 x BC018075 ..

```

```

Align seg 1/1 to: BC018075 from: 1 to: 1873

```

```

1824 G|YTRP.GluArgMetPheGlyThrTyrPheArgValGlyPheTyrGlyT 1840
||||| : : |||||||
1170 GGATGTAAGCGGAGTGTGGACCTATTTTCGTTGGTTTATGAA 1219
|||||
1840 hrlYspheGlyAspLeuAspGluGlnGluPheValTyrLysGluProAla 1856
|||||
1220 CCAAGTTGGGGATTGGATGACACAGAAATTTGTTACAGAGACCTTCA 1269
|||||
1857 lIerhrlYsLeuAlaGluIleSerHisArgLeuGluGlyPheTyrGly 1873
|||||
1270 ATAACCAAACTTGCAGAGATATCTCACAGATTGGAGGATTTTACGAGA 1319
|||||
1873 uArgPheGlyGluAspValValGluValIleLysAspSerSnpProVala 1890
|||||
1320 AAGATTGGAGAGGATGTGGAGTGAAGTATCAAGACTCTAATCCTGTAG 1369
|||||
1890 sPLysCysLysLysLeuAspProAsnLysAlaTyrIleGlnIleThrTyrVal 1906
|||||
1370 ACAATGTAANTAGATCTTAACAAGCATATATTCACATTAACCATGTG 1419
|||||
1907 GluProTyrPheAspThrTyrGluMetLysAspArgIleThrTyrPheAs 1923
|||||
1420 GAGCATACTTGTGACACATATGATGAAGACAGAAATCACCCTTATTCGA 1469
|||||
1923 pLysAsnTyrAsnLeuArgArgPheMetTyrCysThrProPheThrLeuA 1940
|||||
1470 CAAAATTAACAATCTTCGTCAATCATGTACTGTACACCCCTTACTTAG 1519
|||||
1940 sPGLYARAlaHisGlyLysLysGluGlnHisGluGlnPheLysArgLysThrIle 1956
|||||
1520 ATGGCCGCGCCCATGGGAGACTCATGACAAATTCAAAAGGAAGACCAT 1569
|||||
1957 LeuThrThrSerHisAlaPheProTyrIleLysThrArgValAsnValTh 1973
|||||
1570 CTGACTACGCTCATGCTTCTCTATATTAACAAGGTCATATGTCCAC 1619
|||||
1973 rHisLysGluGluIleIleLeuThrProIleGluValAlaIleGluAspM 1990
|||||
1620 TCATTAAGAAGAGATCATCTTAACACCAATGAAAGTCTATGAGGACA 1669
|||||
1990 eGclnLysLysThrGlnGluLeuAlaPheAlaThrHisGlnAspProAla 2006
|||||
1670 TGCAGAAAAGACACAGAGGTGGATTTGCACACATCAGATCCCGCA 1719
|||||
2007 AspProLysMetLeuGlnMetValLeuGlnLysSerValGlyThrThVa 2023
|||||
1720 GACCCCAAAATGCTTCAGATGTACTCCAGGATCTGTAGGACACACAGT 1769
|||||
2023 lAsnGlnGlyProLeuGluValAlaGlnValPheLeuSerGluIleProS 2040
|||||
1770 GAATCAGGGGCTTTGGAAGTTCACAGGTTTCTGTGTAATAACTCA 1819
|||||

```

```

2040 eAspProLysLeuPheArgHisHisAsnLysLeu 2051
|||||
1820 GTGACCAAAAGCTCTTCAGACATCATATAATAACTG 1854
|||||

```

```

seq_name: gb_hlc:BC018076

```

```

seq_documentation_block:

```

```

LOCUS      BC018076      1873 bp      mRNA      linear      HTC      06-DEC-2001
DEFINITION Homo sapiens, clone IMAGE:4792507, mRNA.
ACCESSION  BC018076
VERSION    BC018076.1 GI:17390164
KEYWORDS   HTC.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE  1 (bases 1 to 1873)

```

```

AUTHORS    Strausberg, R.
TITLE       Direct Submission
JOURNAL     Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA

```

```

REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Sequencing Group at the Stanford Human
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcdick@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

```

COMMENT

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 32 Row: b Column: 20
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
 This clone has the following problem: frame shifted.

FEATURES

Location/Qualifiers

1..1873

source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4792507"
/tissue_type="Brain, hypothalamus"
/clone_id="NIH_MGC_96"
/lab_host="DH10B"
/notes="Vector: pBluescript"

```

```

BASE COUNT      604 a      317 c      365 g      587 t
ORIGIN

```

```

alignment_scores:
  Quality: 1199.00      Length: 229
  Ratio: 5.259          Gaps: 1
Percent Similarity: 99.563      Percent Identity: 99.127

```

```

alignment_block:
US-09-737-246-2 x BC018076 ..

```

```

Align seg 1/1 to: BC018076 from: 1 to: 1873

```

```

1824 G|YTRP.GluArgMetPheGlyThrTyrPheArgValGlyPheTyrGlyT 1840
||||| : : |||||||
1170 GGATGTAAGCGGAGTGTGGACCTATTTTCGTTGGTTTATGAA 1219
|||||
1840 hrlYspheGlyAspLeuAspGluGlnGluPheValTyrLysGluProAla 1856
|||||

```

1220 CCAAGTTGGGGATTGGATGACAGAAATTTGTTACAGAGACCCTGCA 1269
1857 lIethrLysLeuAlaGluIleSerHisArgLeuGluGlyPheTyrglyGI 1873
1270 ATAACCAAACTTGCAGAGATATCTCACAGATTGGAGGATTTTACGAGA 1319
1873 uArgPheGlyGluAspValIleGluValIleLysAspSerSerProVala 1890
1320 AAGATTGGACAGAGATGGTGGTAAGTAAATCAAGACTCTAATCCTGTAG 1369
1890 sPLysCysLysLeuAspProAsnLysAlaTyrlleGlnIleIethrTyVal 1906
1370 ACAATGTAAATTAGATCTTACAAAGCATATATTCAGATTACCTATGTG 1419
1907 GluProTyPheAspPThrTyTyrGluMetLysAspArgIleThrTyPheAs 1923
1420 GAGCATCTCTTGGACACATATGAGTGAAGAGACAAATCACCTATTGCA 1469
1923 pLysAsnTyAsnLeuArgArgPheMetTyTyrCysThrProPheThrLeu 1940
1470 CAAAATTTACAAATCTTGGTCATTCATGTAACACCCCTTACTTGTAG 1519
1940 spGIArgAlaHisGlyGluLeuHisGluGlnPheLysArgLysThrIle 1956
1520 ATGGCCGTGCCCATGGGGAATTCATGACAAATTCAAAAGAACACCATT 1569
1957 LeuThrThrSerHisAlaPheProTyrlleLysThrArgValAsnValTh 1973
1570 CTGACTACGCTCATGCTTCTCTTAATATTAACAAAGGGTCATGTGCAC 1619
1973 rHisGlyGluGluIleIleLeuThrProIleGluValAlaIleLyspm 1990
1620 TCATTAAGAAAGATCATCTTAACACCAATTGAAGTTCATTTAGACA 1669
1990 etGlnLysLysThrGlnGluLeuAlaPheAlaThrHisGlnAspProAla 2006
1670 TGCAGAAAAACACACAGAGATGGCATTTGCACACATCAGAGATCCGCA 1719
2007 AspProLysMetLeuGlnMetValLeuGlnLysSerValGlyThrThra 2023
1720 GACCCCAAAATGCTTCAGATGTACTCCAGAGATCTGAGCACACACAGT 1769
2023 lAsnGlnLysProLeuGluValAlaGlnValPheLeuSerGluLeuPro 2040
1770 GAATAGGGGCTTGGAAAGTTGCCAGGTTTCTCTGAAATACCTA 1819
2040 etAspProLysLeuPheArgHisHisAsnLysLeu 2051
1820 GTGACCCAAAGCTCTTCAGACATCATATAAATG 1854
seq_name: gb_est2:BI689126
seq_documentation_block: 869 bp mRNA linear EST 18-snp-2001
LOCUS BI689126 60315930f1 NC_1 CGAP_Mam6 Mus musculus cDNA clone IMAGE:5355798 5',
DEFINITION mRNA sequence.
ACCESSION BI689126 GI:15651755
VERSION BI689126.1 GI:15651755
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
1 (bases 1 to 869)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILMIL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILML at:
<http://image.lnl.gov>
Plate: ILML1904 row: 1 column: 07
High quality sequence stop: 761.
Location/Qualifiers
1. 869
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5355798"
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1: SalI;
Site: 2: NotI; Cloned unidirectionally. Primer: oligo dt.
Library constructed by Life Technologies, Inc.
providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT 223 a 213 c 237 g 196 t
ORIGIN
alignment_scores:
Quality: 1193.00 Length: 271
Ratio: 4.553 Gaps: 4
Percent Similarity: 96.679 Percent Identity: 92.989
alignment_block:
US-09-737-246-2 x BI689126 ..
Align seg 1/1 to: BI689126 from: 1 to: 869
1284 MetIleSerGlnThrValAlaMetAlaIleAlaGlyThrSerValProGI 1300
12 GATATAGCCAGACGAGTTCATGCAATTCAGAGAACATCCGCTCTCA 61
1300 nLeuThrArgProGlySerPheLeuLeuThrSerThrSerGlyArgGlnH 1317
62 GCTGACACAGACGTCAGTTTCTCTCTCAGTCACAGATGCGCGGAC 111
1317 isThrThrPheSerAlaGluSerSerArgSerLeuLeuIleCysLeuLeu 1333
112 ACACCACTTCTCAGCAGAGTCAGACGAGAGCCCTTGTGATCTGTGCTT 161
1334 TrpValLeuLysAsnAlaAspGluThrValLeuGlnLysTrpPheThrAs 1350
162 TGGGTTCTCAAGATGACAGATGAACGGTTCTGCAGAAATGGTTTACGA 211
1350 pLeuSerValLeuGlnLeuAsnArgLeuLeuAspLeuLeuTyrlleCysV 1367
212 TCTCTCAGTCTCTCAGCTGACCGGTTGTATGATCTGCTTACCTGTG 261
1367 aLserCysPheGluTyrlLysGlyLysLys.ValPheGluArgMetAsnSe 1383
262 TATCTTGCTTTGAGTCAAAAGAAAAAACGTAATTTGAACGAATGAATAG 311
1383 rLeuThrPheLysLysSerLysAspMetArgAlaLysLeuGluGluAlaI 1400
312 TTTAACCTTTTAAAGAAATCAAAAGACTGAGACCAAGCTTGAAGAGCA 361
1400 lLeuGlnLysSerIleGlyAlaArgGlnGluMetValArgArgSerArgly 1416
362 TTCTGGGAGCATCGTCCAGGCAAGAAATGTACGCGGAGCGGAGG 411
1417 GlnLeuGluArgSerProSerGlySerAlaPheGlySerGlnGluAsnLe 1433
412 CAGCTCGAGAGAGACCATCTGGAAGCGCTTGGAGAGCAAGAAACCT 461
1433 uArgTrp.ArgLysAspMetThrHisTrpArgGlnAsnThrGluLysLeu 1449
462 GCGGTGAGCAAAAGACATGACTGCGCTCAAGAACTCAAGAAACTG 511

1450 AsplysSerArgAlaGluIleGluHisGluAlaLeuIleAspGlyAsnIle 1466
|||||
512 GATTAAGTCAGAGCAGACATAGAAACAGACACATGATTGAAACCT 561
|||||
1466 uAlaThrGluAlaAsnLeuIleIleLeuAspThrLeuGluIleValG 1483
|||||
562 GCCTACTAGAACCAACCTCATCTAGACACCTAGACATCATCTGCC 611
|||||
1483 InThrValSerValThrGluSerIleLeuGlyGlyValLeu 1499
|||||
612 AGACCTCTCTGTAAAGATCCAAAGAGATCTCGGTGGTGCTA 661
|||||
1500 LysValIleLeuHisSerMetAlaCysAsnGlnSerAlaValTyrLeuG 1516
|||||
662 AAAGTCGCTACAGACATGCGCTGCACCAAGT.GCAGTCTATCTGCA 710
|||||
1516 nHisCysPheAlaThrGlnArgAlaLeuValSerLysPheProGlu.Leu 1532
|||||
711 GCAGTCTCTCGCCACGACAGACCTGCTCAAGTCTCGGAGCTCTT 760
|||||
1533 LeuPheGluGluGluThrGluGlnCysAla.AspleuGlySerLeuIle 1549
:::|||||
761 GTTCGAGGAGACAGACGACGATGTGCGGATCTGTCCCTCCGCTTC 810
|||||
1549 euArg 1550
811 TCCGG 815
seq_name: gb_est2:B6872239

seq_documentation_block: 745 bp mRNA linear EST 29-MAY-2001
LOCUS B6872239
DEFINITION 602790848F1 NCL_CGAP_SG2 Mus musculus cDNA clone IMAGE:4922339 5',
mRNA sequence.
ACCESSION B6872239
VERSION B6872239.1 GI:14222779
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 745)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
http://image.lnl.gov
Plate: LHAM10841 row: O column: 12
High quality sequence stop: 737.
Location/Qualifiers
1..745
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4922339"
/clone_id="NCL_CGAP_SG2"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1;
Not; Site: 2; Salt; Cloned undirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCL_CGAP Library."
BASE COUNT 219 a 167 c 196 g 163 t
ORIGIN

alignment_scores:
Quality: 1146.00 Length: 247
Ratio: 4.736 Gaps: 2
Percent Similarity: 97.976 Percent Identity: 95.142

alignment_block:
US-09-737-246-2 x B6872239 ..

Align seg 1/1 to: B6872239 from: 1 to: 745

1375 LysLysValPheGluArgMetAsnSerLeuThrPheLysSerIleLysAs 1391
|||||
3 AAAAAGTATTGACGAATGAAATAGTTAACTTTAGCAAAATACACAA 52
1391 pMetArgAlaLysLeuGluGluAlaIleLeuGlySerIleGlyAlaArgG 1408
|||||
53 CATGAGAGCCAGCTTGAAGAACCATTTCTGGGAGACATCGTCCAGGC 102
1408 InGluMetValArgArgSerArgGlyGlnLeuGluArgSerProSerGly 1424
|||||
103 AGGAATGTGTACGGGGAGCCGAGGGCAGCTCGAGAGAGACCATCTGGA 152
1425 SerAlaPheGlySerGlnGluAsnLeuArgTTPArgLysAspMetThrH 1441
|||||
153 AGCCCTTTGGGAGCCAAAGAAACCTGCGGTGGAGAAAGACATGACTCA 202
1441 sTPArgGlnAsnThrGluLysLeuAspLysSerArgAlaGluIleGluH 1458
|||||
203 CTGGCGTCAGAACTCAGAGAACTCGATTAAGTCAAGAGCAGAAATGAAAC 252
1458 IsGluAlaLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeuIleIle 1474
|||||
253 ACAGAGCACTGATGATGAAACCTGGCTACTGACAGCAAACTCATCATC 302
1475 LeuAspThrLeuGluIleValGlnThrValSerValThrGluSerLys 1491
|||||
303 TTAGACACGCTAGAGATCATCTCCACAGCTGTCTGTACGGAATCCAA 352
1491 sGluSerIleLeuGlyGlyValLeuLysValLeuLeuHisSerMetAlaC 1508
|||||
353 AGAGAGTATCTGGGTGTGTCTTAAAGTCTGCTACAGACATGGCCT 402
1508 LysAsnGlnSerAlaValTyrLeuGlnHisCysPheAlaThrGlnArgAla 1524
|||||
403 GCAACCAAGTGCAGTCTATCTGCAGCATGCTCCGACGACAGAGGCC 452
1525 LeuValSerLysPheProGluLeuLeuPheGluGluIleThrGluGln 1541
|||||
453 CTGCTCTCAAAAGTTCTCTGAGCTCTTGTAGAGAAAGACAGACAGTGC 502
1541 sAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSerIleGlyT 1558
|||||
503 TCGGATCTGTGCTCCGGCTTCTCCCTCATTTGCAGTAGACATCATGTA 552
1558 hrIleArgSerHisProSerAlaSerLeuTyrLeuLeuMetArgGlnAsn 1574
|||||
553 CAATCCGGTGCACGCTAGTGCCTCCCTTACCTCATGAGAGCAGAAC 602
1575 PheGluIleGlyAsnAsnPheAlaArgValLysMetGlnValProMetSe 1591
|||||
603 TTTGAGATCGGGAATTAACCTTGCAGAGTGAATAATGCAATGCGAGTGC 652
1591 rLeuSerSerLeuValGlyThrSerGlnAsnPheAsn.GluGluPheLeu 1607
|||||
653 ACTGTCTCTTAGTGGGTAGCTCTCAGAATTTTAATCGAAGATTTCTTC 702
1608 Arg.ArgSerLeuLysThrIleLeuThrTyrAlaGlu 1619
|||||
703 AGACCGCTCTTAATAAATAATTGACATATGCTGAA 739
seq_name: gb_hlc:AK013336
seq_documentation_block:

Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAAGACGATCCAGACGCTCTTTTGTTCCTTTTTTTCVN 3']. cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 7.5 and subtraction to
Rot = 37.5. Second strand cDNA was prepared with the primer adapter
of sequence [5'
GAGAGAGACATTCTCGAGTTAATTAAATTAAATCCCCCCCCC 3']. cDNA was cleaved
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOLR.

FEATURES

SOURCE

location/Qualifiers

1..636

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="MCD:MGI:1896206"

/db_xref="taxon:10090"

/clone="2810453G12"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="10, 11 days embryo"

179..>636

/note="data source:SPFR, source key:O9HIQ3, evidence:ISS
homolog to BA16SF24.1.1 (NOVEL PROTEIN SIMILAR TO
DROSOPHILA CG6630 AND CG11376, KIAA1058, RAT TRG (ISOFORM
1)) (FRAGMENT)
putative"

/codon_start=1

/protein_id="BAB28798.1"

/db_xref="GI:12850637"

/translation="MKDRITFPKKNYNLRFMYCTPEPLDGRAHGELHEQFKRTILT
TSASFPIKTRVNTHKETLIIIPLEIAEDMDKKIOELAFAFHODPADPKMLQMVLQ
GSVTYNNGPRLVAOVFELSEIGDPKRLFHNHKKLRICEFDFTKREDALR"

BASE COUNT 202 a 141 c 133 g 160 t

ORIGIN

alignment_scores:

Quality: 1106.00 Length: 211

Ratio: 5.242 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 98.578

Alignment_block:

US-09-737-246-2 x AK013336 ..

Align seg 1/1 to: AK013336 from: 1 to: 636

1856 AAttEtHrLyLeuAlaGlUtlEsErHsArGbEuGlUGlyPhyTrGl 1872

|||||

2 GCAAATAACCMAACTCGCAGAGATTTCACAGATTGGAGGGATTTTATGG 51

1872 yGUatrphegiyGUaspaIvaIgIwAlIlelysAspaAgIIEthTyph 1889

|||||

52 AGAAAGATTTGAGAGAGATGTCCTTGAAGTATCAAGCACTCTATCCTG 101

1889 aAspLyScysLyLeuAspProasnLySaIaTyrtIeGlnIIEthTytr 1905

|||||

102 TAGACAAGCGCAAATTAGATCCAAACAGGCATTAATTACAGTTACCTAC 151

1906 VAlglUpPrOtyrPheaSpThrTytrGlUwEllysaPaAgIIEthTyph 1922

|||||

152 GTGGAGCCCTTCTTGACACCTATGAGTAGAAGGACAGATCACCTATTT 201

1922 eAspLySaInTyraSnLeuArgArgPheMetLyrcTyshPrOpheHrL 1939

|||||

202 TGACAAAATTTATATCTCCGGCGTTTCATGATGACACACCTTCACCT 251

1939 euAspGLyArgAlaHisGlyLUleuHISgLUglnPheLyArgLyshThr 1955

|||||

252 TAGATGGCCGTGCTCATGGGAGCTTCGAACAATTCAACGGGAAGACC 301

1956 lIleutrrHnSerHsAlaPhePrOtyrIlleLyshThrArgValaSnVa 1972

|||||

302 ATCTGACACGCTCATGCTTCCTTACATTAACACAGATCAATGT 351
1972 lTthrHstysglugluIlelleleuThrProIleGluValAlaIleGluA 1989
|||||
352 CACTCACAAGAGAGATATCTTAACCAATGAGTTCTTCGAG 401
1989 sPmetGlnLysLysThrGlnGluLeuAlaPheAlaThrHisGlnAspPro 2005
|||||
402 ACATGCAGAAAAAGACATCAGGAATTTGCATTGGCAACATCAGAGATCCA 451
2006 AlaAspProLysMetLeuGlnMetValLeuGlnGlySerValGlyThrTh 2022
|||||
452 GCAGACCCCAAAATGCTTCAATGGCTCTCCAGGGGCTGTAGAGAACAC 501
2022 rValAsnGlnGlyProLeuGluValAlaGlnValPheLeuSerGluIleP 2039
|||||
502 TGTAAACACAGGACCTCTGAGAGTTGCCAGGTCTTCTGTCTGAATAC 551
2039 rSerAspProLysLeuPheArgHisHisAsnLysLeuArgLeuGlyPhe 2055
|||||
552 CTGTGACCCGAAAGCTTTCAGACATCACAAAGCTCGCCTTTGTTTC 601
2056 LysAspPheThrLysArgCysGlnAspAlaLeu 2066
|||||
602 AAGGACTTACTTAAAGGTGTGAGGATGCTTGA 634
seq_name: gb_est2:BI770146
seq_documentation_block:
LOCUS BI770146 796 bp mRNA linear EST 25-SEP-2001
DEFINITION 603053332F1 NIH_MGC_122 Homo sapiens CDNA clone IMAGE:5203063 5',
mRNA sequence.
ACCESSION BI770146
KEYWORDS BI770146.1 GI:15761724
EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 796)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaups-femail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
plate: LHAM1508 row: p column: 08
High quality sequence stop: 793.
Location/Qualifiers
1..796
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5203063"
/clone_lib="NIH_MGC_122"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: PCMV-SPORT6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH MGC Library."
BASE COUNT 188 a 228 c 231 g 148 t 1 others
ORIGIN

alignment_scores:
Quality: 1104.00 Length: 250
Ratio: 4.738 Gaps: 1
Percent Similarity: 93.200 Percent Identity: 82.800
alignment_block:
US-09-737-246-2 x BI770146 ..
Align seg 1/1 to: BI770146 from: 1 to: 796
1836 GlyPheThrGlyThrLysPheGlyAspLeuAspGlnGluPheValIy 1852
1 GGCCTTACGGCCGCCACCTTGGCTGACCTGGATGACAGAGATTGTCTTA 50
1852 rLysGluProAlaIleThrLysLeuAlaGluIleSerHisArgLeuGlu 1869
|||||
51 CAAGAGCCATCGATCGATCGAAGCTGGCAGAGATCTCACACCGCTGAGG 100
1869 LysPheArgGlyIleArgPheGlyGluAspValAlaGluAlaIleLysAsp 1885
|||||
101 AGTTCTACAGGAGAGATTGGGAGAGAGCTGCTGAGATATCAAGAG 150
1886 SerAsnProValAspLysCysLysLeuAspProAsnLysAlaIleGlu 1902
|||||
151 TCTAACCTGTGGACAAAGTCCAAAGCTTGACTCACAAAGGCCCTACATCA 200
1902 nIleThrTyrValGluProTyrPheAspThrTyrGluMetLysAspArgI 1919
|||||
201 GATACGATATGTGAACCGTACTTTGATACCTACGAGCTCAAGACCGG 250
1919 lThrTyrPheAspLysAsnTyrAsnLeuArgArgPheMetTyrCysThr 1935
|||||
251 TGACCTACTTTGAACCGCACTAGGGCTTGCACATTCCTGTTCTGACAG 300
1936 ProPheThrLeuAspGlyArgAlaHisGlyGluLeuHisGluGlnPhe 1952
301 CCGTTTACGGCCGATGGGCGGCACACGGGAGCTGCCACCAACACA 350
1952 sArgLysThrIleLeuThrTrpSerHisAlaPheProTyrIleLysTha 1969
|||||
351 GCGTAAGACGCTGCTCAGACACGACGCGCTTCCCTTACATTAAGATC 400
1969 rGluAsnValThrHisLysGluGluIlelleuThrProIleGluVal 1985
|||||
401 GCATCCGTGTGTGCCACCGGAGAGAGACGGTGCAGACGCCAGTGAAGTG 450
1986 AlaIleGluAspMetGlnLysLysThrGlnGluLeuAlaPheAlaThrH 2002
451 GCCATCGAGGACATGCAGAAAGAACACAGGGAGCTGCCCTTGCACCGGA 500
2002 sGlnAspProLysAspProLysMetLeuGlnMetValLeuGlnGlySerV 2019
|||||
501 GCAGAGCCACACAGATGCTAAGATGCTACAGATGTCTTCAGAGGCTCTG 550
2019 aLysGlnThrValAsnGlnGlyProLeuGluValAlaGlnValAlaPheLeu 2035
|||||
551 TAGGGCCACACCGTGACACAGGCTCCCTGGAGGTGGCCAGGCTTTTGA 600
2036 SerGluIleProSerAspProLysLeuPheArgHisHisAsnLysLeuA 2052
|||||
601 GCAGAGATCCCGGAAGACCCCAAGCTTCTCCGGCATTCACAAATTTGCG 650
2052 gLeuGlyPheLysAspPheThrLysArgCysGlnAspAlaLeuArgLys. 2068
|||||
651 GCTTGCTTCAAGACCTTCTCCAAAGATGTGAGAGATGCGTGGGAAAC 700
2069 AsnLysSerLeuIleGlyProValGlnGlyLysGluIleArgGluLeu 2084
|||||
701 AATAAGGCCCTGATTTGGCCGCGACACAGAGAGTACACCGTGAAGCTG 748
seq_name: gb_est2:BF309712

seq_documentation_block:
LOCUS BF309712 898 bp mRNA linear EST 21-NOV-2000
DEFINITION 601891854F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4137732 5',
mRNA sequence.
ACCESSION BF309712
VERSION BF309712 GI:11257158
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 898)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM1045 row: k column: 13
high quality sequence stop: 732.
Location/Qualifiers
1..898
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:4137732"
/clone_lib="NIH_MGC_17"
/tissue_type="rhabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: muscle, Vector: pOTB7, Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(C). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 276 a 182 c 228 g 212 t
ORIGIN
alignment_scores:
Quality: 1092.00 Length: 282
Ratio: 4.233 Gaps: 4
Percent Similarity: 91.489 Percent Identity: 87.234
alignment_block:
US-09-737-246-2 x BF309712 ..
Align seg 1/1 to: BF309712 from: 1 to: 898
1378 phegluargmetasnserleuthrphelyslyserlyaspmetargal 1394
|||||
3 tttagacgaatgattacgttgacctttaagaatcaaaagacatgagacg 52
1394 alysleugluguallaleuenglyserllleglyalaarglnglunety 1411
|||||
53 aaaccttgaaagaaacgtattcttgggacacatagtgcccagacgaagaaatgg 102
1411 alatgargserargglyglnglengluargserproserglyseralaph 1427
|||||
103 ttagcgccagacccgagacagctcgaggaagcccatctggaaagtgaccttt 152
1428 glyserlngluasnleuargtrparglyaspmetthrhistrpargl 1444
|||||
153 ggaagtcagaaatattgaggtgaggaagaaatatactacatggcgctca 202
1444 nasnthglulysleuasplyssetargalagluillegluhisglualal 1461

|||||
203 aaacacagaaacgttgacaaatcaagacagacagatttgacacgaagcac 252
1461 eutleaspglyasnleualaathrglualaasnleuilelleuaspthr 1477
|||||
253 tgattgatgaaacctggctacagaaacaaacctaatcatattttagataca 302
1478 leugluillevalvalglnthrvalservalthrghlserlysglunserl 1494
|||||
303 tttagaagattgtt.gttcagaccgtttctgtgacgaattccaaagagacgt 351
1494 eleuglyglvalleuvalysvalleuuehissermetalaacyasnglns 1511
|||||
352 tcttgatggagctgcttaaaagtgcgtactacacagcattgctgaaccaa 401
1511 eralavaltryleuglnhisgysphealathrglnargalaleuvalser 1527
|||||
402 gtccagtttatctacaaacactggtttgctaacacagacccttggtttca 451
1528 lyspheprogluueuuephegluglugluthrghlucngysalasphe 1544
|||||
452 aaatttctgaaccttattttgaagaagacagacagactgctgatttt 501
1544 ucysleuargleuarghisgyserserierllleglythrilearg 1561
502 atgcctcagacctctccagacagctgtagcagatcgctacaaatagcgt 551
1561 erhisproseralaserleutryrleuueuemetarglnasnphegluile 1577
|||||
552 cacaccgccagctgctccctttacctactaatgagcgaataactt.gagatc 600
1578 glyasnasnphelaargvalylsmetglnvalprometserleuserse 1594
601 gggaatgaactttgccaggctttaaatagcaggtaacaaatgctatgcctc 649
1594 rleuvalglythrserglnasnphasnlglnupheleuargargserl 1611
650 .ctggggcgacacatctcaggaatttttaagagaatctttaaagacgtcttc 699
1611 eulysthrileleuthr.tyralaeglunaspheugluueargluuth 1627
|||||
700 ttagactattattgacactatgctctggagaaatctggcaatggcgggag 749
1627 rthrpeproaspglnvalgln.aspleuvalpheasnleuhismetile 1643
|||||
750 caccttgccggatcaagccagagatcgsgtt.....cacaccact 790
1644 leuseraspthrvalylsmetlysglnhisglnluaspptro 1657
|||||
791 atgacgtttctgctacgtctgaga.....gaggaacccc 823

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2002, 08:33:13 ; Search time 646.39 Seconds

(without alignments)
18136.252 Million cell updates/sec

Title: US-09-737-246-1

Perfect score: 6828

Sequence: 1 gtcgcgcgtccgcgcagcagc.....ctcgtgcgcgcgcagcgcctc 6828

Scoring table:

IDENTITY_NTC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 segs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_032802:*

1: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6828	100.0	6828	22	AAHA1934 Human CLASP-3 CDNA
2	4123	60.4	4143	22	AAHA1911 Preliminary human
3	4123	60.4	4144	22	AAHA1912 Human CLASP-3 CDNA
4	4123	60.4	4200	22	AAHA1953 CLASP-3 nucleotide
5	3044.8	44.6	3173	22	AAHA1952 CLASP-3 nucleotide
6	2132.8	31.2	6372	22	AAHA3851 Human CLASP-7 enco
7	1749	25.6	7215	22	AA507373 Human cDNA encodin
8	1079	15.8	4027	22	AA507382 Human cDNA associat
9	919.6	13.5	2148	22	AAHA3850 Preliminary human

10	841.6	12.3	2014	22	AA507381
C 11	93.89	9.1	93.89	22	ABL17736
12	565	8.3	1652	22	AAH99572
13	459.8	6.7	2427	22	AAI61331
14	434.4	6.4	2437	22	AAI59545
15	370	5.4	1834	24	AA562625
16	334.4	4.9	933	22	AAH14933
17	324.6	4.8	1566	22	AAH17340
18	321.2	4.7	5589	23	ABL17737
19	304.8	4.5	2610	22	AA532647
20	304.8	4.5	6454	22	AA508334
21	301.6	4.4	3472	22	AA508358
22	298.4	4.4	4393	22	AA508357
C 23	281	4.1	365	14	AA060222
24	245.8	3.6	1273	21	AA509301
25	244.4	3.6	4108	22	AAH18725
26	216.2	3.2	260	22	AAH19419
27	214.6	3.1	3899	22	AA532675
28	207.8	3.0	5688	21	AA5087969
29	207.8	3.0	5688	21	AAH14825
30	207.8	3.0	7277	24	AA518951
31	196.4	2.9	1605	22	AA507383
32	190.4	2.8	996	23	AA587322
33	190.4	2.8	2036	22	AAH14086
34	183.4	2.7	5214	21	AA5087968
35	183.4	2.7	5214	21	AAH14824
36	180.2	2.6	441	22	AAH184302
37	178.2	2.6	220	22	AAH1937
38	174.6	2.6	586	21	AA507563
39	171.6	2.5	2165	21	AA507649
40	168	2.5	211	22	AAH14945
41	167.4	2.5	221	22	AAH14949
42	154.6	2.3	201	22	AAH14943
43	153.6	2.2	6816	21	AA507524
44	153.6	2.2	7506	22	AAH19118
45	148	2.2	981	23	AA5070846

ALIGNMENTS

RESULT 1	
AAHA1934	AAHA1934 standard; cDNA; 6828 BP.
ID	AAHA1934
XX	XX
AC	AAHA1934:
XX	XX
DT	30-AUG-2001 (first entry)
XX	XX
DE	Human CLASP-3 CDNA sequence SEQ ID NO:1.
XX	XX
KW	Human: cadherin-like asymmetry protein; CLASP; CLASP-3; immune response; cell surface molecule; transmembrane protein; immunosuppressive; vaccine; antiinflammatory; antitachytic; antianaemic; dermatological; uropathic; K W
KW	ophthalmological; antihypertensive; nephrotoxic; antithyroid; antidiabetic; K W
KW	neuroprotective; antistimatic; antibacterial; antisense therapy; K W
XX	XX
OS	Homo sapiens.
XX	XX
PN	WO200142297-A2.
XX	XX
PD	14-JUN-2001.
XX	XX
PF	13-DEC-2000; 2000MO-US34171.
XX	XX
PR	13-DEC-1999; 99US-0170453.
PR	14-JAN-2000; 2000US-0176195.
PR	14-FEB-2000; 2000US-0182296.
PR	11-APR-2000; 2000US-0196267.
PR	11-APR-2000; 2000US-0196460.
PR	11-APR-2000; 2000US-0196527.
PR	11-APR-2000; 2000US-0196528.

Human DNA associat
Drosophila melanog
Human protein enco
Human polynucleoti
Human polynucleoti
CDNA sequence #412
Human CLASP-3 geno
Human cDNA sequenc
Drosophila melanog
Human cDNA encodin
Human cDNA encodin
Human cDNA encodin
Human cDNA encodin
Human brain Expres
Human secreted pro
Human nervous syst
Human CLASP-3 geno
Human cDNA encodin
Human cDNA encodin
Human CLASP relate
Human cDNA encodin
Human DNA associat
DNA encoding novel
Human cDNA sequenc
Mouse CLASP-1 nucl
DNA encoding a mur
Human polynucleoti
Human CLASP-3 geno
Human ORF2504
Human CLASP-3 geno
Human CLASP-3 geno
Human CLASP-3 geno
Human ORF79 p
Angiogenesis assoc
DNA encoding novel

CC The present invention describes an isolated polypeptide (I) comprising a
CC nucleotide (nt) sequence (Sta) that has at least 90% identity to the
CC sequence given in AAH1934 and is immunologically cross-reactive with
CC the derived amino acid (aa) sequence (Sib) given in AAB9495 or shares a
CC biological function with native CLASP-3. (I) has immunosuppressive,
CC anti-inflammatory, antiarthritic, antinaeemic, dermatological, uterohelptic,
CC antihistamological, antihypertensive, nephrotoxic, antihypertrophic, antidiabetic,
CC neuroprotective, antistomatitic and antibacterial activities, and can be
CC used in antisense therapy, vaccine production and gene therapy. CLASP-3
CC related sequences can be used in preventing or treating a CLASP-3
CC mediated disease, preferably an autoimmune disease by inhibiting an
CC immune response. The autoimmune disease is caused or exacerbated by
CC increased activity of T helper 1 lymphocytes (Th1). AAH1933 to AAH1953
CC and AAB9491 to AAB99507 represent sequences which are used in the
CC exemplification of the present invention. CLASP-3 is localised in the
CC chromosome location 1p31.1.
CC
CC
CC Sequence 6828 BP; 2082 A; 1384 C; 1432 G; 1930 T; 0 other;
CC
CC

[illegible]

QY	481	ccaaacaagtttctgtaactgataagctccagaatgagcaacagctaccagaatgataca	540
Db	481	ccaaaaaagatttctgtaactgataagcctccagatgagcaacagctaccagaatgataca	540
QY	541	gattgacctaaagaagcgttcaatgtaataagatgatacccccaagggttagcgggcgt	600
Db	541	gattgacctaaagaagcgttcaatgtaataagatgatacccccaagggttagcgggcgt	600
QY	601	agtatccttgacttgaaanaatctcactcctgatagtcttgcttcccaattactgataca	660
Db	601	agtatccttgacttgaaanaatctcactcctgatagtcttgcttcccaattactgataca	660
QY	661	actccaatgagaatataagacgcgtcaagatgatagacccaagaagatccaacgltacaa	720
Db	661	actccaatgagaatataagacgcgtcaagatgatagacccaagaagatccaacgltacaa	720
QY	721	gaacttttgcttgatcaccacccagatgaggaagaaccaatgaaggcttagtgt	780
Db	721	gaacttttgcttgatcaccacccagatgaggaagaaccaatgaaggcttagtgt	780
QY	781	ccgtgatacccaagaacatttgtgcgaagaactcttgtaaaatgactatacactcaag	840
Db	781	ccgtgatacccaagaacatttgtgcgaagaactcttgtaaaatgactatacactcaag	840
QY	841	tttgaatatgaaatgaaacccattttgtcaagtttggtcttataatgatagtaagaaaag	900
Db	841	tttgaatatgaaatgaaacccattttgtcaagtttggtcttataatgatagtaagaaaag	900
QY	901	aaaaagattcagaanaacttatatttgaccttaattctgagagagtgaaaggttgtata	960
Db	901	aaaaagattcagaanaacttatatttgaccttaattctgagagagtgaaaggttgtata	960
QY	961	cgtaacacatgacacacgtgctgcacattaccactctgcaagaatcacaacattttctatc	1020
Db	961	cgtaacacatgacacacgtgctgcacattaccactctgcaagaatcacaacattttctatc	1020
QY	1021	acttatccttcccaagatgttttctgtataaagctgaaaaaagctctacagaaga	1080
Db	1021	acttatccttcccaagatgttttctgtataaagctgaaaaaagctctacagaaga	1080
QY	1081	gacatttgagaggtgtgcagaacatatataatttcaagaagcagatgtccccaagat	1140
Db	1081	gacatttgagaggtgtgcagaacatatataatttcaagaagcagatgtccccaagat	1140
QY	1141	aaagaanaaacgtgagaaactgaaagatgcaagcagatcagtttgcgaagaagcttgagaa	1200
Db	1141	aaagaanaaacgtgagaaactgaaagatgcaagcagatcagtttgcgaagaagcttgagaa	1200
QY	1201	tatgcatagcttttgcgttgactgcaatccatttaagataatgtttagcagtgctgg	1260
Db	1201	tatgcatagcttttgcgttgactgcaatccatttaagataatgtttagcagtgctgg	1260
QY	1261	agtttggaagaagattctacaagaatagaatctagtaactgagaaagcaaaaggtcttgg	1320
Db	1261	agtttggaagaagattctacaagaatagaatctagtaactgagaaagcaaaaggtcttgg	1320
QY	1321	tcagagaagaagaattctatgattgttgcgacagctacacttgaaagaacaaagvtgga	1380
Db	1321	tcagagaagaagaattctatgattgttgcgacagctacacttgaaagaacaaagvtgga	1380
QY	1381	gatgatagtctgtaacttgagagacttgcgacagctactctcacagtgacaatattttt	1440
Db	1381	gatgatagtctgtaacttgagagacttgcgacagctactctcacagtgacaatattttt	1440
QY	1441	aagcagaagaagacgcgttaagtatgataagatctctacaatctccgtgtgataagaa	1500
Db	1441	aagcagaagaagacgcgttaagtatgataagatctctacaatctccgtgtgataagaa	1500
QY	1501	aggccatcttctgtactagcggcagcaaaacatttaagcgtcagctccaagatagat	1560
Db	1501	aggccatcttctgtactagcggcagcaaaacatttaagcgtcagctccaagatagat	1560
QY	1561	tctccgcaccttgaaatctcccatattgtccttaactccgcgaactgtcttcaagtgaagctt	1620

|||||
Db 1561 tctccgcacccgtgaaaaatccccaattatctgctaactccgcagctgcgtccaaagtgaagctt 1620
OY 1621 taacctgacgtagaagcttagaacctaccagaagaattcttagagttcccgcaaggagatt 1680
Db 1621 taacctgacgtagaagcttagaacctaccagaagaattcttagagttcccgcaaggagatt 1680
OY 1681 tatgttccaacaacatacttcacagaatcttctctacatacataccctcagaatcttaattt 1740
Db 1681 tatgttccaacaacatacttcacagaatcttctctacatacataccctcagaatcttaattt 1740
OY 1741 gccaatcgtaaggcttctgcttagaanaataacagtgaaagttcaagttatgtttgagag 1800
Db 1741 gccaatcgtaaggcttctgcttagaanaataacagtgaaagttcaagttatgtttgagag 1800
OY 1801 gatccaagcaatgctgcgcggtgaatcttctgttaaatctgacgtgctgaagaatttcaag 1860
Db 1801 gatccaagcaatgctgcgcggtgaatcttctgttaaatctgacgtgctgaagaatttcaag 1860
OY 1861 gaagcctatacagccgttagatatacatcaacaagctctcctgattcttcatagagaatcaag 1920
Db 1861 gaagcctatacagccgttagatatacatcaacaagctctcctgattcttcatagagaatcaag 1920
OY 1921 gttaaagcttcctgctactttaactgacatcaatcactgcttcttaacttttcatatggt 1980
Db 1921 gttaaagcttcctgctactttaactgacatcaatcactgcttcttaacttttcatatggt 1980
OY 1981 agttgtcaacaanaaacaataactcctctgtgaaacaccagtttgatataatgatacca 2040
Db 1981 agttgtcaacaanaaacaataactcctctgtgaaacaccagtttgatataatgatacca 2040
OY 2041 atgcttcagaatggacggttgaagaactgacgttctgtcctgcgcagctcatctggaanaa 2100
Db 2041 atgcttcagaatggacggttgaagaactgacgttctgtcctgcgcagctcatctggaanaa 2100
OY 2101 ccaccacaaggcttatctctgactgtccctcgtgaaggttccctacactgycagaaatggtta 2160
Db 2101 ccaccacaaggcttatctctgactgtccctcgtgaaggttccctacactgycagaaatggtta 2160
OY 2161 gataatcaacaaggctgttttaaatggtgaagttgtgtcttccgctatccataaccaca 2220
Db 2161 gataatcaacaaggctgttttaaatggtgaagttgtgtcttccgctatccataaccaca 2220
OY 2221 gatccttatctctgcaaaatttttgcctcgtgcaatgctctgatagtaacaccgttccca 2280
Db 2221 gatccttatctctgcaaaatttttgcctcgtgcaatgctctgatagtaacaccgttccca 2280
OY 2281 gtccgaattgggacatgycgaatcatatggaataaatacttagaanaatgaaatggaagcagt 2340
Db 2281 gtccgaattgggacatgycgaatcatatggaataaatacttagaanaatgaaatggaagcagt 2340
OY 2341 attccagaacatgaaatcatccacagctggaacccagttgtccgatttctctctctgcta 2400
Db 2341 attccagaacatgaaatcatccacagctggaacccagttgtccgatttctctctctgcta 2400
OY 2401 gataaacatgatacttttagttatataagactcctcgtgtcaatgtcgtgccaataagttaacta 2460
Db 2401 gataaacatgatacttttagttatataagactcctcgtgtcaatgtcgtgccaataagttaacta 2460
OY 2461 ggtcaagatcttttgaagcatalgycataatataaatactgacttcaaaaaacttggaa 2520
Db 2461 ggtcaagatcttttgaagcatalgycataatataaatactgacttcaaaaaacttggaa 2520
OY 2521 ggaatatatgacagcatgycagagaacagcctcttgacatataatacttatgttttc 2580
Db 2521 ggaatatatgacagcatgycagagaacagcctcttgacatataatacttatgttttc 2580
OY 2581 cgcttacacaataacttaacctaatatcatcacacaggttcctgggggtttgggaagatca 2640
Db 2581 cgcttacacaataacttaacctaatatcatcacacaggttcctgggggtttgggaagatca 2640
OY 2641 gtgcattatgcacaatgcttagactctgagggtgagaccctgcaagccttaattaaatcgt 2700
Db 2641 gtgcattatgcacaatgcttagactctgagggtgagaccctgcaagccttaattaaatcgt 2700

Db 2641 gtgcattatgccaacatggtctagatctgcgtgagacactgcaagccttaattaaatcgt 2700
OY 2701 tctcgaagccttagtaataagcaatccagatatatctgggactcccaagcttcaaccaatgat 2760
Db 2701 tctcgaagccttagtaataagcaatccagatatatctgggactcccaagcttcaaccaatgat 2760
OY 2761 gaagtgcatacatcatcgggagtaaggcttttagatctgcctcaatcttcgtgttlaacact 2820
Db 2761 gaagtgcatacatcatcgggagtaaggcttttagatctgcctcaatcttcgtgttlaacact 2820
OY 2821 ggtgtgtccaaaagctgtcccatgtgggataccaaacccagttccaagtgtcaaatcaacag 2880
Db 2821 ggtgtgtccaaaagctgtcccatgtgggataccaaacccagttccaagtgtcaaatcaacag 2880
OY 2881 gctatgagtcgaagttgtgatatcgtatgtcttcgtgcaacagaagctgaagtttcttaaca 2940
Db 2881 gctatgagtcgaagttgtgatatcgtatgtcttcgtgcaacagaagctgaagtttcttaaca 2940
OY 2941 acattaaacgggacgcttaacaaactaaagaacttttccagaggagctgtcttgacgtg 3000
Db 2941 acattaaacgggacgcttaacaaactaaagaacttttccagaggagctgtcttgacgtg 3000
OY 3001 gttgtgttcagtgagcaggttccgggaatcagcttgcgaacaaagctgttctttttgaa 3060
Db 3001 gttgtgttcagtgagcaggttccgggaatcagcttgcgaacaaagctgttctttttgaa 3060
OY 3061 ttaatgttaagaagcatggtgcacattatacttaattgaataaactggaagcttccagag 3120
Db 3061 ttaatgttaagaagcatggtgcacattatacttaattgaataaactggaagcttccagag 3120
OY 3121 aaaagtcgttttcocgaagaagtttcatatgatatgacatgtgacgtctgttcaagacgtatgct 3180
Db 3121 aaaagtcgttttcocgaagaagtttcatatgatatgacatgtgacgtctgttcaagacgtatgct 3180
OY 3181 agtgcataatgttcaacagattcaagaagacacagaagaatggttgaagactcaatacaagc 3240
Db 3181 agtgcataatgttcaacagattcaagaagacacagaagaatggttgaagactcaatacaagc 3240
OY 3241 ctgcacattcttctcaaatgatctgttctgttctatgggacagaagattgttttaagcctt 3300
Db 3241 ctgcacattcttctcaaatgatctgttctgttctatgggacagaagattgttttaagcctt 3300
OY 3301 ataaagcttcctgtatataaaagaagtgcttcttcaaaagcttctcaatccgaatccagtgct 3360
Db 3301 ataaagcttcctgtatataaaagaagtgcttcttcaaaagcttctcaatccgaatccagtgct 3360
OY 3361 ctggtgtccttgaagctgtgaatttcttaacgaatcatctgcagatcagacatagtltaca 3420
Db 3361 ctggtgtccttgaagctgtgaatttcttaacgaatcatctgcagatcagacatagtltaca 3420
OY 3421 ttaaaacttccctgcagacttaacttaactccacgtgcatccatcaactctgtttctctc 3480
Db 3421 ttaaaacttccctgcagacttaacttaactccacgtgcatccatcaactctgtttctctc 3480
OY 3481 gcaacatctcaagatcttgaatttcttaacgaatgtacagaaagcaaaaagattgtcaaatatg 3540
Db 3481 gcaacatctcaagatcttgaatttcttaacgaatgtacagaaagcaaaaagattgtcaaatatg 3540
OY 3541 ttgtgaattatcgtgtccttcccgcaacagcaatatttggcagaagctgtgttacaagag 3600
Db 3541 ttgtgaattatcgtgtccttcccgcaacagcaatatttggcagaagctgtgttacaagag 3600
OY 3601 ctggcgtgatacttttagaccctgatgtctgaagacatggttggatgtcatagaagaatgatc 3660
Db 3601 ctggcgtgatacttttagaccctgatgtctgaagacatggttggatgtcatagaagaatgatc 3660
OY 3661 aatatgttacacaatattatcttccagttcaacagactcagaaccccggttaacttgaacctag 3720
Db 3661 aatatgttacacaatattatcttccagttcaacagactcagaaccccggttaacttgaacctag 3720
OY 3721 ataaagctcagatggtgcacatgttgtatatacctctgatgtgtattatcatatgaaactgt 3780
Db 3721 ataaagctcagatggtgcacatgttgtatatacctctgatgtgtattatcatatgaaactgt 3780

QY 3781 cctcagctgatagtatttaccagaaactcacatcaacgaggaagaccatttgtatagcc 3840
|||||
Db 3781 cctcagctgatagtatttaccagaaactcacatcaacgaggaagaccatttgtatagcc 3840
QY 3841 actgatactatgaaagttagagcggaagtatgataagccagaccgttgcataggaacac 3900
|||||
Db 3841 actgatactatgaaagttagagcggaagtatgataagccagaccgttgcataggaacac 3900
QY 3901 gcaaggagatactggtccctccaactaacaagcgctggcagtttccctccacgtcaacagat 3960
|||||
Db 3901 gcaaggagatactggtccctccaactaacaagcgctggcagtttccctccacgtcaacagat 3960
QY 3961 ggcagggcaacacactaccttcttcagcagaatccaagtccaagcctttgatactgctact 4020
|||||
Db 3961 ggcagggcaacacactaccttcttcagcagaatccaagtccaagcctttgatactgctact 4020
QY 4021 tgggtctcaaaaaatgacagatgaaacagttctacagaagtgggtttacagatctctcagtc 4080
|||||
Db 4021 tgggtctcaaaaaatgacagatgaaacagttctacagaagtgggtttacagatctctcagtc 4080
QY 4081 ttgcagctcaaacccggtctttagactctgtcttatactctgtgtgtcttgcattgagataaa 4140
|||||
Db 4081 ttgcagctcaaacccggtctttagactctgtcttatactctgtgtgtcttgcattgagataaa 4140
QY 4141 gggaaaaaagtgtttgaaacgatatgacttgaccccttaagaaatcaaaaagacatgaga 4200
|||||
Db 4141 gggaaaaaagtgtttgaaacgatatgacttgaccccttaagaaatcaaaaagacatgaga 4200
QY 4201 gcaaaagcttgaaagaagctactctctggagacataagtgccagcaagaatgylgacggcga 4260
|||||
Db 4201 gcaaaagcttgaaagaagctactctctggagacataagtgccagcaagaatgylgacggcga 4260
QY 4261 agccgaggaacagctctcgagagaagaccacatctggaagtgcctttggaatccaagaatattg 4320
|||||
Db 4261 agccgaggaacagctctcgagagaagaccacatctggaagtgcctttggaatccaagaatattg 4320
QY 4321 aggttgaaagaagatatgactcactgcgtccaacaacagagaagccttgacaatacaga 4380
|||||
Db 4321 aggttgaaagaagatatgactcactgcgtccaacaacagagaagccttgacaatacaga 4380
QY 4381 gcaagagatlgacaacgaaacgacatgattgattgaaacctggtctacagaagcaaacctatc 4440
|||||
Db 4381 gcaagagatlgacaacgaaacgacatgattgattgaaacctggtctacagaagcaaacctatc 4440
QY 4441 atttaagatacttagagattgtgttccagccgttctctgaaacggatcccaagaagac 4500
|||||
Db 4441 atttaagatacttagagattgtgttccagccgttctctgaaacggatcccaagaagac 4500
QY 4501 attcttgtagagtgctaaagtgctactacacagcaltgqccctgtaaccaaagtgcagtc 4560
|||||
Db 4501 attcttgtagagtgctaaagtgctactacacagcaltgqccctgtaaccaaagtgcagtc 4560
QY 4561 tatctacaacacgtcttctgtacacagagagccttgglttcaaaagtctcctgtaacctta 4620
|||||
Db 4561 tatctacaacacgtcttctgtacacagagagccttgglttcaaaagtctcctgtaacctta 4620
QY 4621 ttgtaagaagaagacagagcagtggtgtgatttatagcttcagcgtctcccgacatctgagc 4680
|||||
Db 4621 ttgtaagaagaagacagagcagtggtgtgatttatagcttcagcgtctcccgacatctgagc 4680
QY 4681 agtagcactcggtatacaatacaggtccacccccagtgctccctcttaccactaaagagcga 4740
|||||
Db 4681 agtagcactcggtatacaatacaggtccacccccagtgctccctcttaccactaaagagcga 4740
QY 4741 aactcttgagatlgggaataactcttgccaagggttaaaatgacagtgtaaccaatgtcactac 4800
|||||
Db 4741 aactcttgagatlgggaataactcttgccaagggttaaaatgacagtgtaaccaatgtcactac 4800
QY 4801 tcccttgtaggagacatctccagaattttaagaagaattcttaagacgttctctaaagact 4860
|||||
Db 4801 tcccttgtaggagacatctccagaattttaagaagaattcttaagacgttctctaaagact 4860

QY 4861 atattgacatattgtcgaagaagatctgtaattgagggaaacaacatttccctgatcagtc 4920
|||||
Db 4861 atattgacatattgtcgaagaagatctgtaattgagggaaacaacatttccctgatcagtc 4920
QY 4921 caagatctgglttccaatccatcatgattcttctctgatactgtgaaaaatgaaagaaac 4980
|||||
Db 4921 caagatctgglttccaatccatcatgattcttctctgatactgtgaaaaatgaaagaaac 4980
QY 4981 caaggaagatccctgaaatgtgtattgtactaaagttaacagatggcaagggttaacagac 5040
|||||
Db 4981 caaggaagatccctgaaatgtgtattgtactaaagttaacagatggcaagggttaacagac 5040
QY 5041 tctccagagcgaatttgacccggttctgacgaacatgacggcgaaagcactccaacaacgaacat 5100
|||||
Db 5041 tctccagagcgaatttgacccggttctgacgaacatgacggcgaaagcactccaacaacgaacat 5100
QY 5101 catgctgaaagctgacagatgtctagttccacatcagcagcactgtgtgtaataatttgagc 5160
|||||
Db 5101 catgctgaaagctgacagatgtctagttccacatcagcagcactgtgtgtaataatttgagc 5160
QY 5161 atgctggaagcagggaataatctctctgtggagatgtgttaactttcagaataattccact 5220
|||||
Db 5161 atgctggaagcagggaataatctctctgtggagatgtgttaactttcagaataattccact 5220
QY 5221 aatgttttagaagaatctcggtctcagaatgatagtgtatctccagatgaaagaagtatc 5280
|||||
Db 5221 aatgttttagaagaatctcggtctcagaatgatagtgtatctccagatgaaagaagtatc 5280
QY 5281 tgcctcgaaanaacttacttactagttcaggaactgttggatctcgtgaaacaagcagctgc 5340
|||||
Db 5281 tgcctcgaaanaacttacttactagttcaggaactgttggatctcgtgaaacaagcagctgc 5340
QY 5341 tccctctctatgctgtgcattgataagacagtttaagaagtttaacaagttacttctct 5400
|||||
Db 5341 tccctctctatgctgtgcattgataagacagtttaagaagtttaacaagttacttctct 5400
QY 5401 attcatgaaagctaatcgggatacgcaagaacatccacatcatcgtataaacttcaagaa 5460
|||||
Db 5401 attcatgaaagctaatcgggatacgcaagaacatccacatcatcgtataaacttcaagaa 5460
QY 5461 gcaatcaagcaaaatgttcatcaagatctggtcgaggagcgaatgttggcaactattc 5520
|||||
Db 5461 gcaatcaagcaaaatgttcatcaagatctggtcgaggagcgaatgttggcaactattc 5520
QY 5521 cgtgtgtgttttatgaaaccaaagtcgggatttggatgaaacgaagatttgtttacaag 5580
|||||
Db 5521 cgtgtgtgttttatgaaaccaaagtcgggatttggatgaaacgaagatttgtttacaag 5580
QY 5581 gaacctgcaataaccaaacttgcagaagatctcacaagatlgagaagattttaacgagaa 5640
|||||
Db 5581 gaacctgcaataaccaaacttgcagaagatctcacaagatlgagaagattttaacgagaa 5640
QY 5641 agatttgagagatgttggatgaagtaatacaagacttaactcctgttagacaagtgtaaa 5700
|||||
Db 5641 agatttgagagatgttggatgaagtaatacaagacttaactcctgttagacaagtgtaaa 5700
QY 5701 tttagatcccaaaaaggacatatattagattaccttctgtggaacaaacttttagacatat 5760
|||||
Db 5701 tttagatcccaaaaaggacatatattagattaccttctgtggaacaaacttttagacatat 5760
QY 5761 gaagataaagacagaatccacacttgcagaaanaattaaaccttcgttcogattcaglac 5820
|||||
Db 5761 gaagataaagacagaatccacacttgcagaaanaattaaaccttcgttcogattcaglac 5820
QY 5821 tgtacaaccttaactttagatggcggtgccatgggggaactctcatgaaacatccaagaag 5880
|||||
Db 5821 tgtacaaccttaactttagatggcggtgccatgggggaactctcatgaaacatccaagaag 5880
QY 5881 aagacacatctgactagcgtctatgcttcccttatataaacaagggttaagtgcact 5940
|||||
Db 5881 aagacacatctgactagcgtctatgcttcccttatataaacaagggttaagtgcact 5940
QY 5941 cataaagaagatcatctttaacacaatttgaaagtgtctattgagagacatgcaagaaag 6000
|||||

|||||
Db 5941 cataaagaagatcatcttaccaccaatltgaatgtctatgtgagcatgcagaaaaag 6000
Oy 6001 acacaggaggtggcattgtgcaacatcagatgccgagaccccaaatgtctcagaatg 6060
Db 6001 acacaggaggtggcattgtgcaacatcagatgccgagaccccaaatgtctcagaatg 6060
Oy 6061 gtactccaggagctgttagcaccacagtgatcaggggctcttggaagtgtgccaggtt 6120
Db 6061 gtactccaggagctgttagcaccacagtgatcaggggctcttggaagtgtgccaggtt 6120
Oy 6121 ttctgtctggaataactcagtgacccaagctcttcagacatcataaataactgtgacatc 6180
Db 6121 ttctgtctggaataactcagtgacccaagctcttcagacatcataaataactgtgacatc 6180
Oy 6181 tgccttaagaatttactcaaaagggtgtgaagatgccttaaaaaataaagaagtttaatt 6240
Db 6181 tgccttaagaatttactcaaaagggtgtgaagatgccttaaaaaataaagaagtttaatt 6240
Oy 6241 gggccggttcaaaaggagatcacaagggaatttggggaacatcttcgctctaagaagagc 6300
Db 6241 gggccggttcaaaaggagatcacaagggaatttggggaacatcttcgctctaagaagagc 6300
Oy 6301 ccttaagccctcagatcaagaagaagtcctcagttatcccaagccagttatgtcttccct 6360
Db 6301 ccttaagccctcagatcaagaagaagtcctcagttatcccaagccagttatgtcttccct 6360
Oy 6361 gccacagagatccttcagtcagatggaatgttcgcaaatgtgacttaaatcaatga 6420
Db 6361 gccacagagatccttcagtcagatggaatgttcgcaaatgtgacttaaatcaatga 6420
Oy 6421 ctgtttatcatctgcgaagaagccatgtatcaacatcagtggtgaagaagcatatg 6480
Db 6421 ctgtttatcatctgcgaagaagccatgtatcaacatcagtggtgaagaagcatatg 6480
Oy 6481 gaacacaacatggaatggaatctcgtgaatattatcattcattggaagaatgctggccaag 6540
Db 6481 gaacacaacatggaatggaatctcgtgaatattatcattcattggaagaatgctggccaag 6540
Oy 6541 aaaatcaaatgtgaatgtgttaacgcttgaagatcagtcagtcagtcagtcagtcagtc 6600
Db 6541 aaaatcaaatgtgaatgtgttaacgcttgaagatcagtcagtcagtcagtcagtcagtc 6600
Oy 6601 gtaacaagctgtatcctttaaagacatttaagtactcaaaaggtaacatactaac 6660
Db 6601 gtaacaagctgtatcctttaaagacatttaagtactcaaaaggtaacatactaac 6660
Oy 6661 attatttaccatagcctaagaagttaaaaattatcattcattcattcattcattcattcatt 6720
Db 6661 attatttaccatagcctaagaagttaaaaattatcattcattcattcattcattcattcatt 6720
Oy 6721 tatacacacattatagatcattcttgaccatttaaatgtagtaagtatttttaa 6780
Db 6721 tatacacacattatagatcattcttgaccatttaaatgtagtaagtatttttaa 6780
Oy 6781 aggtactaaaaaatatgtgaatgtttactcgtgcgcgcagggcgctc 6828
Db 6781 aggtactaaaaaatatgtgaatgtttactcgtgcgcgcagggcgctc 6828

RESULT 2
AAH41911
ID AAH41911 standard; cDNA: 4143 BP.

AAH41911;

30-AUG-2001 (first entry)

Preliminary human CLASP-3 cDNA sequence Fig 1.

Human; cadherin-like asymmetry protein; CLASP; CLASP-3; immune response;
cell surface molecule; transmembrane protein; immunosuppressive; vaccine;
antiinflammatory; antiatheritic; antihaemic; dermatological; uropathic;

KW ophthalmological; antineumatic; nephrotropic; antithyroid; antidiabetic;
KW neuroprotective; antiasthmatic; antibacterial; antisense therapy;
KW gene therapy; chromosome 1p31.1; ss.
OS Homo sapiens.
PN WO200142297-A2.
PD 14-JUN-2001.
PE 13-DEC-2000; 2000MO-US34171.
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
PR 11-APR-2000; 2000US-0196267.
PR 11-APR-2000; 2000US-0196460.
PR 11-APR-2000; 2000US-0196527.
PR 11-APR-2000; 2000US-0196528.
PR 11-APR-2000; 2000US-0547276.
PR 13-OCT-2000; 2000US-0240503.
PR 13-OCT-2000; 2000US-0240508.
XX (ARBO-) ARBOR VITA CORP.
XX Lu P, Garman JD, Candia AF;
XX WPI; 2001-375003/39.
XX P-PSDB; AAB99493.
PT Novel isolated cadherin-like asymmetry protein (CLASP) -3, useful for
PT useful in preventing or treating a CLASP-3-mediated disease e.g.
PT autoimmune disease -
XX
XX Example 5; Fig 1; 189pp; English.
XX
XX The present invention describes an isolated polypeptide (I) comprising a
XX nucleotide (nt) sequence (Sia) that has at least 90% identity to the
XX sequence given in AAH4194 and is immunologically cross-reactive with
XX the derived amino acid (aa) sequence (Sib) given in AAB99495 or shares a
XX biological function with native CLASP-3. (I) has immunosuppressive,
XX antiinflammatory, antiatheritic, antianaemic, dermatological, uropathic,
XX ophthalmological, antineumatic, nephrotropic, antithyroid, antidiabetic,
XX neuroprotective, antiasthmatic and antibacterial activities, and can be
XX used in antisense therapy, vaccine production and gene therapy. CLASP-3
XX related sequences can be used in preventing or treating a CLASP-3
XX mediated disease, preferably an autoimmune disease by inhibiting an
XX immune response. The autoimmune disease is caused or exacerbated by
XX CC and AAB99491 to AAB99507 represent sequences which are used in the
XX CC exemplification of the present invention. CLASP-3 is localised in the
XX chromosome location 1p31.1.
XX
SQ Sequence 4143 BP; 1251 A; 829 C; 890 G; 1173 T; 0 other;

Query Match 60.4%; Score 4123; DB 22; Length 4143;
Best Local Similarity 99.9%; Pred No. 0;
Matches 4136; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Oy 2693 taatcgtctcgaagccttagatagcaatccagatatactggaatcccaagtcac 2752
Db 5 taatcgtctcgaagccttagatagcaatccagatatactggaatcccaagtcac 64

Oy 2753 cagatgatgaagtcgatacatatcgggaaggaaggtttagatcgctccaattcctgg 2812
Db 65 cagatgatgaagtcgatacatatcgggaaggaaggtttagatcgctccaattcctgg 124

Oy 2813 ttacacgtgtgtccaaagctgcccagtggaatcccaacccagttccaagtgaagat 2872
Db 125 ttacacgtgtgtccaaagctgcccagtggaatcccaacccagttccaagtgaagat 184

Oy 2873 caacacaggtatgatacgaaagttgataatgtcttcgcacacagagacgtcaagtt 2932

Db 105 caacacagctatgatctgaagtgttaatcgtatgtcttcgcacacagagcgtcaagtt 244
QY 2933 tcttaacaacatlaacggagcgttaaccaactaaaaagcttttccacgagagctgctt 2992
Db 245 tcttcaacaacatlaacggagcgttaccactaaaaagcttttccacgagagctgctt 304
QY 2993 tgcagttggtttgttcgagttgcagcgttcgggaaatcgagctttgcaaaaagcctgtctt 3052
Db 305 tgcagttggtttgttcgagttgcagcgttcgggaaatcgagctttgcaaaaagcctgtctt 364
QY 3053 tttttggaattaaatggtlaaagagcgtggtgcaccatttaacttaattatataacttggg 3112
Db 365 tttttggaattaaatggtlaaagagcgtggtgcaccatttaacttaattatataacttggg 424
QY 3113 ctccaagaaaaaagctcgctttccagaagcttcaatgataatgacgtcttgcagca 3172
Db 425 ctccaagaaaaaagctcgctttccagaagcttcaatgataatgacgtcttgcagca 484
QY 3173 cgattgctagtgataatgtttcacgatttcagaaagcacagagaatggtttgagagactca 3232
Db 485 cgattgctagtgataatgtttcacgatttcagaaagcacagagaatggtttgagagactca 544
QY 3233 atacaagccttgcattcttcttcaatga tctgttgcgttatgtgacagagagatttgtt 3292
Db 545 atacaagccttgcattcttcttcaatga tctgttgcgttatgtgacagagagatttgtt 604
QY 3293 ttaagcctataaagctccgcgtcataaacaagtggtcttcaaaagcttaccattacccaagtc 3352
Db 605 ttaagcctataaagctccgcgtcataaacaagtggtcttcaaaagcttaccattacccaagtc 664
QY 3353 ccaggtctctgtgtgccttcgagagctggaattcttcaagaaatcctgcagtcattgagcact 3412
Db 665 ccaggtctctgtgtgccttcgagagctggaattcttcaagaaatcctgcagtcattgagcact 724
QY 3413 atgttaacatlaaacttaaccctgcagcttaacttaactcaaccctgcacatccacacactcttg 3472
Db 725 atgttaacatlaaacttaaccctgcagcttaacttaactcaaccctgcacatccacacactcttg 784
QY 3473 ttctcttcgcaacatctcagagttcttggaattttctcagaagtacagaaccaaaagattg 3552
Db 785 ttctcttcgcaacatctcagagttcttggaattttctcagaagtacagaaccaaaagattg 844
QY 3533 caaatatgtttgaattatccgtgcgtcttcgcgcaaacagcatattttgcagagacttgtt 3592
Db 845 caaatatgtttgaattatccgtgcgtcttcgcgcaaacagcatattttgcagagacttgtt 904
QY 3593 taacagagctggtcgtlcatltttagaccctgagtgcgaagagactgttttgatgtcataaga 3652
Db 905 taacagagctggtcgtlcatltttagaccctgagtgcgaagagactgttttgatgtcataaga 964
QY 3653 aagttacaatatgtgttaacaatattactctcagtcacagacttcgagaccggtgactctg 3712
Db 965 aagttacaatatgtgttaacaatattactctcagtcacagacttcgagaccggtgactctg 1024
QY 3713 accctcagataaaggctcgagttgcgcatgtgtatctacccctgataatgtatatacag 3772
Db 1025 accctcagataaaggctcgagttgcgcatgtgtatctacccctgataatgtatatacag 1084
QY 3773 aaactgttaacctgaactgtatgtatlttacagaaactcacatcaacgaggaaccaattt 3832
Db 1085 aaactgttaacctgaactgtatgtatlttacagaaactcacatcaacgaggaaccaattt 1144
QY 3833 gtaattgcaactgtatgtatgtaaagttagagagcggaagatgataagacagacggttgca 3892
Db 1145 gtaattgcaactgtatgtatgtaaagttagagagcggaagatgataagacagacggttgca 1204
QY 3893 tggcaatcgcaaggacatcggttccctcaactaaacaagcctgcgacgttcttccctcaagt 3952
Db 1205 tggcaatcgcaaggacatcggttccctcaactaaacaagcctgcgacgttcttccctcaagt 1264
QY 3953 caacgagtgcgaggaacaacactaccttttaagcagaaatcaagtgcagccttttgatct 4012
|||||

Db 1265 caacgagttgcgaggaacacacactaccttttcagcagaaatcaagtcgaagccttttgatct 1324
QY 4013 gtctactttgtgttctcraaaaatgcaagatgaaacagttcttaagaagtgtttacagatc 4072
Db 1325 gtctactttgtgttctcraaaaatgcaagatgaaacagttcttaagaagtgtttacagatc 1384
QY 4073 ttcagcttcgtgcagcttaaacccgtctatagatctgttttaactcctgtgtctgtcttg 4132
Db 1385 ttcagcttcgtgcagcttaaacccgtctatagatctgttttaactcctgtgtctgtcttg 1444
QY 4133 agtataaaggaaaaaagtgttttgaacgaatgaatagcttgacctttaagaatcaaaag 4192
Db 1445 agtataaaggaaaaaagtgttttgaacgaatgaatagcttgacctttaagaatcaaaag 1504
QY 4193 acatgaaagcaaaagcttgaagaagctatcttcttgaggcataggtgcgaagcaagaattg 4252
Db 1505 acatgaaagcaaaagcttgaagaagctatcttcttgaggcataggtgcgaagcaagaattg 1564
QY 4253 tacggcgaaagccgaggaagcgtcgaagaagcccatctggaagtgccttgcgaagtcaag 4312
Db 1565 tacggcgaaagccgaggaagcgtcgaagaagcccatctggaagtgccttgcgaagtcaag 1624
QY 4313 aaaatttgaagtgcgaggaagaatataactcaactgcgtgcgaacaaacacagagaagcttgaca 4372
Db 1625 aaaatttgaagtgcgaggaagaatataactcaactgcgtgcgaacaaacacagagaagcttgaca 1684
QY 4373 aatcaagagcagaagatttgaaacagagcctgatttgaaagcctggttacaagaagcga 4432
Db 1685 aatcaagagcagaagatttgaaacagagcctgatttgaaagcctggttacaagaagcga 1744
QY 4433 acctaatcatlttagatatacatagaaatgttttcaagacggttcttcgtaacggaatcca 4492
Db 1745 acctaatcatlttagatatacatagaaatgttttcaagacggttcttcgtaacggaatcca 1804
QY 4493 aagaagacatcttctgtgtgagtgcttaaaagtctactcaacagcatgacctgaaccaaa 4552
Db 1805 aagaagacatcttctgtgtgagtgcttaaaagtctactcaacagcatgacctgaaccaaa 1864
QY 4553 gtgcagtttaactacaacatggttttgcctacaagaagagcctgtgttcaaaagtttccgtg 4612
Db 1865 gtgcagtttaactacaacatggttttgcctacaagaagagcctgtgttcaaaagtttccgtg 1924
QY 4613 aaactctaatctgaagaagagacagagcagtgltgctaatatgatgcctcaagcttctccagac 4672
Db 1925 aaactctaatctgaagaagagacagagcagtgltgctaatatgatgcctcaagcttctccagac 1984
QY 4673 actgttagcagtagcatcgtgtacaataatcgtgtcacaccccagtgctctcttaactactaa 4732
Db 1985 actgttagcagtagcatcgtgtacaataatcgtgtcacaccccagtgctctcttaactactaa 2044
QY 4733 tgaaggcaaaacttgaagattgggaataactttgcaggggttaaaatgcaggttaaccaatgt 4792
Db 2045 tgaaggcaaaacttgaagattgggaataactttgcaggggttaaaatgcaggttaaccaatgt 2104
QY 4793 cactatcctcctgtgttgcgacatctcagaatlttaatgaagaattcttaagaagcttctc 4852
Db 2105 cactatcctcctgtgttgcgacatctcagaatlttaatgaagaattcttaagaagcttctc 2164
QY 4853 taaagatataatgacataatgcgtgaagaagaatcttgaatgtgagggaacaaacatttccg 4912
Db 2165 taaagatataatgacataatgcgtgaagaagaatcttgaatgtgagggaacaaacatttccg 2224
QY 4913 atcaggtccaggaatcaggttttcaactcacaatagtttcttcttga tactgtgaaatga 4972
Db 2225 atcaggtccaggaatcaggttttcaactcacaatagtttcttcttga tactgtgaaatga 2284
QY 4973 aggaacacagaggaatcctcgtgaaatgttgatgtatctaaatgttacaagaatttgcgaaggtt 5032
Db 2285 aggaacacagaggaatcctcgtgaaatgttgatgtatctaaatgttacaagaatttgcgaaggtt 2344
QY 5033 accagaccttccaga----gcgattgaactgtgttcagaagaatgycaggaagcactcag 5089
Db 2345 accagaccttccagaatctgcattgcattgacctgtgtgcagaacatgycaggaagcactcag 2404
|||||

OY 5090 aacgaacatcatgctgaaagctgcacagtgctcagctccacacgaagcagcactgtgtctg 5149
|||||
Db 2405 aacgaacatcatgctgaaagctgcacagtgctcagctccacacgaagcagcactgtgtctg 2464
OY 5150 aatattgagcatctgtagaggaacccggaataatcttctgtgtagtggtgttaacattcaga 5209
|||||
Db 2465 aatattgagcatctgtagaggaacccggaataatcttctgtgtagtggtgttaacattcaga 2524
OY 5210 atatttcacaaatggttttaagaagaatctgcgtctcagatgagtggttatccacagatg 5269
|||||
Db 2525 atatttcacaaatggttttaagaagaatctgcgtctcagatgagtggttatccacagatg 2584
OY 5270 aagaaggtatctgctcgtgaaataacttactgagtcagagactgtgtggattactggac 5329
|||||
Db 2585 aagaaggtatctgctcgtgaaataacttactgagtcagagactgtgtggattactggac 2644
OY 5330 aagcagcgtctctctctcctctatgctgctgcatgtagtagcagcttaagaagtttacaaag 5389
|||||
Db 2645 aagcagcgtctctctctcctctatgctgctgcatgtagtagcagcttaagaagtttacaaag 2704
OY 5390 tactatctcctatcatbtaagaactaatcgggaatgcaagaacataaccacaaattcatgta 5449
|||||
Db 2705 tactatctcctatcatbtaagaactaatcgggaatgcaagaacataaccacaaattcatgta 2764
OY 5450 aacttcaagaagcatctcagcaaaatgtgtcaacagatgactgctgtggagcggatgtttg 5509
|||||
Db 2765 aacttcaagaagcatctcagcaaaatgtgtcaacagatgactgctgtggagcggatgtttg 2824
OY 5510 gcaactatttcctgctgtgtgtttttatggaacaaagttccgggattgtgtagaagaagaat 5569
|||||
Db 2825 gcaactatttcctgctgtgtgtttttatggaacaaagttccgggattgtgtagaagaagaat 2884
OY 5570 ttgtttacaagaagcctgcacaaacaaactgtcagagatactccacagatttggaggat 5629
|||||
Db 2885 ttgtttacaagaagcctgcacaaacaaactgtcagagatactccacagatttggaggat 2944
OY 5630 tttaacggaagaagatttggagagagatgtgtgtgaagtaatcaaaagactctaatcctgtag 5669
|||||
Db 2945 tttaacggaagaagatttggagagagatgtgtgtgaagtaatcaaaagactctaatcctgtag 3004
OY 5690 acaagtgtaaatattgatactcctaacaagcatattcagatattactaagtgtggagccatact 5749
|||||
Db 3005 acaagtgtaaatattgatactcctaacaagcatattcagatattactaagtgtggagccatact 3064
OY 5750 ttgacacatcagatgtagaagacagaatcacctatttcgacaaataattacatcttcgtc 5809
|||||
Db 3065 ttgacacatcagatgtagaagacagaatcacctatttcgacaaataattacatcttcgtc 3124
OY 5810 gattcatgtactgtacaccccttactttagatgtagcgtgtgccatgtaggaaactcatgac 5869
|||||
Db 3125 gattcatgtactgtacaccccttactttagatgtagcgtgtgccatgtaggaaactcatgac 3184
OY 5870 aattcaaaaagaagacacatttgcactacgtctcaatgcttccttactataataaaagg 5929
|||||
Db 3185 aattcaaaaagaagacacatttgcactacgtctcaatgcttccttactataataaaagg 3244
OY 5930 tcaatgtaccatcaaaagaagatcatcttaacaacaaatggaagtgtgtatagagaca 5989
|||||
Db 3245 tcaatgtaccatcaaaagaagatcatcttaacaacaaatggaagtgtgtatagagaca 3304
OY 5990 tgcagaanaagacacagagattgtgcatattgcaacacatcagatcccgacagaccacaaa 6049
|||||
Db 3305 tgcagaanaagacacagagattgtgcatattgcaacacatcagatcccgacagaccacaaa 3364
OY 6050 tgccttcagatgtactccagagatctgtaggacacacagatgaatcaggggccttggag 6109
|||||
Db 3365 tgccttcagatgtactccagagatctgtaggacacacagatgaatcaggggccttggag 3424
OY 6110 ttgcccaggttttctgtcgtcgaataactagtagcccaaaagctttagacatcata 6169
|||||
Db 3425 ttgcccaggttttctgtcgtcgaataactagtagcccaaaagctttagacatcata 3484

OY 6170 aactgcgactctgctttaaagatttactaaaggtgtgtagatgctcttaagaanaata 6229
|||||
Db 3485 aactgcgactctgctttaaagatttactaaaggtgtgtagatgctcttaagaanaata 3544
OY 6230 agagcttaatttggcgcttcaaaaaggtatcacaaggggaatttgggaaactatttcgc 6289
|||||
Db 3545 agagcttaatttggcgcttcaaaaaggtatcacaaggggaatttgggaaactatttcgc 3604
OY 6290 cttaaaagagccctacagacccatgacacagaagaagctccctcagttatccaagccagat 6349
|||||
Db 3605 cttaaaagagccctacagacccatgacacagaagaagctccctcagttatccaagccagat 3664
OY 6350 gcttgcctcgcacagagatctcctcagtcgaatgagcttgcgaatgtgatactccta 6409
|||||
Db 3665 gcttgcctcgcacagagatctcctcagtcgaatgagcttgcgaatgtgatactccta 3724
OY 6410 aactgaatgcactgtgtttatcatctcgaagaagagccatgatatcaactcgaagtga 6469
|||||
Db 3725 aactgaatgcactgtgtttatcatctcgaagaagagccatgatatcaactcgaagtga 3784
OY 6470 aagatctatggaacacacatggaatggaatctgtgaatattatattcatggaagatg 6529
|||||
Db 3785 aagatctatggaacacacatggaatggaatctgtgaatattatattcatggaagatg 3844
OY 6530 cagtgccaaagaataatcatcaaatgtagatgttgaacgcttggagacatcgtctatgtt 6589
|||||
Db 3845 cagtgccaaagaataatcatcaaatgtagatgttgaacgcttggagacatcgtctatgtt 3904
OY 6590 ctaagtgtgggtaaaagctgtatcttttaagaatatttaagtactaaaggtacact 6649
|||||
Db 3905 ctaagtgtgggtaaaagctgtatcttttaagaatatttaagtactaaaggtacact 3964
OY 6650 atacattacatatttatacatagcgaaggtlaaaaaattatcacttcaattcgt 6709
|||||
Db 3965 atacattacatatttatacatagcgaaggtlaaaaaattatcacttcaattcgt 4024
OY 6710 atttttaattatcacocattatagatcacttatttggaccattttaaattgtagat 6769
|||||
Db 4025 atttttaattatcacocattatagatcacttatttggaccattttaaattgtagat 4084
OY 6770 gcttattttaaaggtactaaaaataatgtgaatgttactcctgtgcgcccaggccctc 6828
|||||
Db 4085 gcttattttaaaggtactaaaaataatgtgaatgttactcctgtgcgcccaggccctc 4143

RESULT 3
AAH41912
ID AAH41912 standard; cDNA: 4144 BP.
XX
AC AAH41912;
XX
DT 30-AUG-2001 (first entry)
XX
DE Human CIASP-3 cDNA sequence Fig 4.
XX
XX Human; cadherin-like asymmetry protein; CIASP; CIASP-3; immune response;
cell surface molecule; transmembrane protein; immunosuppressive; vaccine;
antiinflammatory; antiarthritic; antianemic; dermatological; urothelial;
ophthalmological; antitumoratic; nephrotropic; antithyroid; antidiabetic;
neuroprotective; antilasthmatic; antibacterial; antisense therapy;
KW gene therapy; chromosome 1p31.1; ss.
XX
XX Homo sapiens.
OS
XX
PN WO200142297-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000MO-US34171.
XX
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.

PR 11-APR-2000; 2000US-0196267.
PR 11-APR-2000; 2000US-0196460.
PR 11-APR-2000; 2000US-0196527.
PR 11-APR-2000; 2000US-0196528.
PR 11-APR-2000; 2000US-0547276.
PR 11-APR-2000; 2000US-0547276.
PR 13-OCT-2000; 2000US-0240503.
PR 13-OCT-2000; 2000US-0240508.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu P, Garman JD, Candia AF;
XX WPI; 2001-375003/39.
DR P-PSDB; AAB99494.
XX
XX
PT Novel isolated cadherin-like asymmetry protein (CLASP) -3, useful for
PT useful in preventing or treating a CLASP-3-mediated disease e.g.
PT autoimmune disease
XX
PS Disclosure: Fig 4A; 189pp; English.
XX
XX The present invention describes an isolated polypeptide (I) comprising a
CC nucleotide (nt) sequence (51a) that has at least 90% identity to the
CC sequence given in AAH41934 and is immunologically cross-reactive to the
CC the derived amino acid (aa) sequence (51b) given in AAB99495 or shares a
CC biological function with native CLASP-3. (I) has immunosuppressive,
CC antiinflammatory, antirheumatic, antianemic, dermatological, uropathic,
CC ophthalmological, antihemetic, nephroprotective, antidiabetic,
CC neuroprotective, antitumor and antibacterial activities, and can be
CC used in antisense therapy, vaccine production and gene therapy. CLASP-3
CC related sequences can be used in preventing or treating a CLASP-3
CC mediated disease, preferably an autoimmune disease by inhibiting an
CC immune response. The autoimmune disease is caused or exacerbated by
CC increased activity of T helper 1 lymphocytes (TH1). AAH41893 to AAH41953
CC and AAB99491 to AAB99507 represent sequences which are used in the
CC exemplification of the present invention. CLASP-3 is localised in the
CC chromosome location 1p31.1.
XX
SQ Sequence 4144 BP; 1252 A; 829 C; 890 G; 1173 T; 0 other;

Query Match 60.4%; Score 4123; DB 22; Length 4144;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4136; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 2693 taaatcgttctcgaagccttaagtaagcaatccagatatacttggagctccacgtcac 2752
Db 6 taaatcgttctcgaagccttaagtaagcaatccagatatacttggagctccacgtcac 65
QY 2753 cagatgataagttcgatcaatcatcgcggagtaaggtttgatcgctccaattcctgg 2812
Db 66 cagatgataagttcgatcaatcatcgcggagtaaggtttgatcgctccaattcctgg 125
QY 2813 ttaacactggtgtccaaaagctgcccaatggggaatccaacccaagtcgaagtgcagaat 2872
Db 126 ttaacactggtgtccaaaagctgcccaatggggaatccaacccaagtcgaagtgcagaat 185
QY 2873 caacacaggtatgataaggtttgaatctatgtcttcgcacacagagacgttcaatt 2932
Db 186 caacacaggtatgataaggtttgaatctatgtcttcgcacacagagacgttcaatt 245
QY 2933 tcttcaacaacattaaacggagccttacaactaaataaagcttttcaacagagcgtgct 2992
Db 246 tcttcaacaacattaaacggagccttacaactaaataaagcttttcaacagagcgtgct 305
QY 2993 tgcagttggtgttcttcgagtcgagcgttcgggaatcagccttgcacaacgctgttct 3052
Db 306 tgcagttggtgttcttcgagtcgagcgttcgggaatcagccttgcacaacgctgttct 365
QY 3053 ttttgaattaatgtaagagcagtgacacattatacttataatgataacttggg 3112
Db 366 ttttgaattaatgtaagagcagtgacacattatacttataatgataacttggg 425

QY 3113 ctccaagaaaagtcgttttccagaacggtttcagatgatacttgcagctctgttcagca 3172
Db 426 ctccaagaaaagtcgttttccagaacggtttcagatgatacttgcagctctgttcagca 485
QY 3173 cgaatgtagatagatagtttccagatctcagaagacacagaatggttggagactca 3232
Db 486 cgaatgtagatagatagtttccagatctcagaagacacagaatggttggagactca 545
QY 3233 atacaagccttgatcttctcgaatgattgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3292
Db 546 atacaagccttgatcttctcgaatgattgtgtgtgtgtgtgtgtgtgtgtgtgtgt 605
QY 3293 ttagcctataaagctcgtgataaacaagtggtcttcaaaagcttactactaccgaatc 3352
Db 606 ttagcctataaagctcgtgataaacaagtggtcttcaaaagcttactactaccgaatc 665
QY 3353 ccagttctgtgtgtcttcgttgaagcgtgatttctcagaatcatctgcagatagcact 3412
Db 666 ccagttctgtgtgtcttcgttgaagcgtgatttctcagaatcatctgcagatagcact 725
QY 3413 atgttaattaaactaacctcgaacttacttactcactcactcactcactcactcact 3472
Db 726 atgttaattaaactaacctcgaacttacttactcactcactcactcactcactcact 785
QY 3473 ttcttctgacaatcctcagaagttcgtgatttctcagaatgatacagaacccaagatg 3532
Db 786 ttcttctgacaatcctcagaagttcgtgatttctcagaatgatacagaacccaagatg 845
QY 3533 caaatatgttgaattatcgtgtccttcgcgcacacagcaattattggcagaactgtgt 3592
Db 846 caaatatgttgaattatcgtgtccttcgcgcacacagcaattattggcagaactgtgt 905
QY 3593 ttaacagagcgtgtgtatatttgacacctgtatcgaagacgttttgattgataaaga 3652
Db 906 ttaacagagcgtgtgtatatttgacacctgtatcgaagacgttttgattgataaaga 965
QY 3653 aagtcacaaatagatgatacaaatctactcctcagtcagcactcagaccgcgtactctg 3712
Db 966 aagtcacaaatagatgatacaaatctactcctcagtcagcactcagaccgcgtactctg 1025
QY 3713 accctcagataaagcgtcgtgtgtccatgttgcatactcctcgtatgttatatactg 3772
Db 1026 accctcagataaagcgtcgtgtgtccatgttgcatactcctcgtatgttatatactg 1085
QY 3773 aaactgtacctcagctgtatgattttacagaacactacaactacaacagagaagacaatt 3832
Db 1086 aaactgtacctcagctgtatgattttacagaacactacaactacaacagagaagacaatt 1145
QY 3833 gataagcactgtatgattatgaaagtgaagcggagatgataaagccagaccgttgcca 3892
Db 1146 gataagcactgtatgattatgaaagtgaagcggagatgataaagccagaccgttgcca 1205
QY 3893 tggcaatcgcagagacatcgtgtccctcaactaaacgaagcgtgcaattcctcctaagt 3952
Db 1206 tggcaatcgcagagacatcgtgtccctcaactaaacgaagcgtgcaattcctcctaagt 1265
QY 3953 caacagagtgagcagcaacacactacacttttcagcagaatcaaatgaaacgtttgatct 4012
Db 1266 caacagagtgagcagcaacacactacacttttcagcagaatcaaatgaaacgtttgatct 1325
QY 4013 gtctacttgggtctcacaatatgacagatgaaacattctacagaagtggtttacagalc 4072
Db 1326 gtctacttgggtctcacaatatgacagatgaaacattctacagaagtggtttacagalc 1385
QY 4073 tctcagcttgcagcttaaacggctattagatctgtcttatactctgtgtgtgtgtgtgt 4132
Db 1386 tctcagcttgcagcttaaacggctattagatctgtcttatactctgtgtgtgtgtgtgt 1445
QY 4133 agtataaagggaataaagtgttggacgaatgaaatgagcttgacatttaagaatcaaaag 4192
Db 1446 agtataaagggaataaagtgttggacgaatgaaatgagcttgacatttaagaatcaaaag 1505
QY 4193 acatgagagcaagccttgaaagactattcttggagacatagtgccaggaagaatg 4252

Db 1506 acatgagagaaaagcttggaaagacattcttgggagcaataggctgcagggcaagaatg 1565
QY 4253 tacggcaagcccgaggaacagctcagagaaagcccatctgtgaagtgtccttggaaagcaag 4312
Db 1566 taaggcgaaagcccgaggaacagctcagagaaagcccatctgtgaagtgtccttggaaagcaag 1625
QY 4313 aaaaattgaggctggaggaagaaatatactactcgtgcgttcaaaaacacagaaagccttgaca 4372
Db 1626 aaaaattgaggctggaggaagaaatatactactcgtgcgttcaaaaacacagaaagccttgaca 1685
QY 4373 aatcaagaagcagagatctggaacacgaaagcacatgttatgtgaaaacctggtctacaagaaca 4432
Db 1686 aatcaagaagcagagatctggaacacgaaagcacatgttatgtgaaaacctggtctacaagaaca 1745
QY 4433 acctaatcatcttagatatactatagatgttgtttcagaccgtttttgtaagggatctca 4492
Db 1746 acctaatcatcttagatatactatagatgttgtttcagaccgtttttcgttaacggatctca 1805
QY 4493 aagagagacatctctgtgtgagtgctaaagaatgctactacacagacatggtcgttaaccaca 4552
Db 1806 aagagagacatctctgtgtgagtgctaaagaatgctactacacagacatggtcgttaaccaca 1865
QY 4553 gtgcagttatctacaacacctgtttgtctacacagagagccttggttcaaaagttcctgt 4612
Db 1866 gtgcagttatctacaacacctgtttgtctacacagagagccttggttcaaaagttcctgt 1925
QY 4613 aacctattttgaaagagagacagagacagatgttgtattatctgcccagcttctccag 4672
Db 1926 aacctattttgaaagagagacagagacagatgttgtattatctgcccagcttctccag 1985
QY 4673 actgtagcagtagacatcgtgacaaatacgtgcacacccacagtgctcctcttaactactaa 4732
Db 1986 actgtagcagtagacatcgtgacaaatacgtgcacacccacagtgctcctcttaactactaa 2045
QY 4733 tggaggcaaaactttgagatctgggaataaactcttgccaaggttaaatgccaagtacaatgt 4792
Db 2046 tggaggcaaaactttgagatctgggaataaactcttgccaaggtttaaagtgcagtgtaaccatgt 2105
QY 4793 cactatcctcctgtgtggtgcacatctcagaattttaatgagaagattcttaagaagcttcc 4852
Db 2106 cactatcctcctgtgtggtgcacatctcagaattttaatgagaagattcttaagaagcttcc 2165
QY 4853 taaagacatatactgacatactgtcgtgaagaagatctggaaatgaggaagaacaacttccgt 4912
Db 2166 taaagacatatactgacatactgtcgtgaagaagatctggaaatgaggaagaacaacttccgt 2225
QY 4913 atcaggtccagagatcgtgtttccaactccatatagtatcttctgtaactgtgaaaaatga 4972
Db 2226 atcaggtccagagatcgtgtttccaactccatatagtatcttctgtaactgtgaaaaatga 2285
QY 4973 aggaacaacagagagatcctggaatgtttgatatatagtacaaatgtccaaaggttt 5032
Db 2286 aggaacaacagagagatcctggaatgtttgatatatagtacaaatgtccaaaggttt 2345
QY 5033 accagacctctcaga -- gcaattgacctgtgtgcagaacatgagcgaagccaagcatcag 5089
Db 2346 accagacctctcagaatcgtcgaattgacctgtgtgcagaacatgagcgaagccaagcatcag 2405
QY 5090 aacgaagaacatctgtcgaagcgtgcacagtgcttaagtcacactcagcaacactgtgtgtcgt 5149
Db 2406 aacgaagaacatctgtcgaagcgtgcacagtgcttaagtcacactcagcaacactgtgtgtcgt 2465
QY 5150 aatatattgacatgctggaagcagcgaataatcttctctggtggagtgtgtaaaatttcaga 5209
Db 2466 aatatattgacatgctggaagcagcgaataatcttctctggtggagtgtgtaaaatttcaga 2525
QY 5210 atattcatctaagttttagaagaatctgcggtctcagaatgaatgtgtatctccagatg 5269
Db 2526 atattcatctaagttttagaagaatctgcggtctcagaatgaatgtgtatctccagatg 2585
QY 5270 aagaaggtatctgctcgtggaataactttaactagtcaggaactgtgtggatattactgtgaac 5329

Db 2586 aagaagratctgctctgynaatactttaactgagtcagagacttgtggatctactgtgaac 2645
QY 5330 aagcagctgcttccctctctatagctggtcatgtagaagcaggttaaatgattacaag 5389
Db 2646 aagcagctgcttccctctctatagctggtcatgtagaagcaggttaaatgattacaag 2705
QY 5390 tactattctctatctaaagctaaatcgcggaatgcagaagaactatcccaactatcatgtga 5449
Db 2706 tactattctctatctaaagctaaatcgcggaatgcagaagaactatcccaactatcatgtga 2765
QY 5450 aactccaagaagcatccagcaaaatgttccatcagagatcgtgctggagcggatgttg 5509
Db 2766 aactccaagaagcatccagcaaaatgttccatcagagatcgtgctggagcggatgttg 2825
QY 5510 gcaactattctggtgtgttttagaaccaaattcggggagtttgatgaacaaagat 5569
Db 2826 gcaactattctggtgtgttttagaaccaaattcggggagtttgatgaacaaagat 2885
QY 5570 ttggttacaagagagccttccaataaccaactgtgcagagatactcaacagatctggaggtat 5629
Db 2886 ttggttacaagagagccttccaataaccaactgtgcagagatactcaacagatctggaggtat 2945
QY 5630 ttacgagaaagatctggagagagatgtgtgtgaagtaatcaaaagactctaaatcctgtag 5689
Db 2946 ttacgagaaagatctggagagagatgtgtgtgaagtaatcaaaagactctaaatcctgtag 3005
QY 5690 acaagtgtaaattagatctcaacaaagcatalatccaattacatctatgttgagccatct 5749
Db 3006 acaagtgtaaattagatctcaacaaagcatalatccaattacatctatgttgagccatct 3065
QY 5750 ttgacacatatgagaatgaagacaaatccactatcttcgaaanaaatatacatatctctgc 5809
Db 3066 ttgacacatatgagaatgaagacaaatccactatcttcgaaanaaatatacatatctctgc 3125
QY 5810 gattcatgtactgtacaccccttacttatagatggtcgtgcccagtggtggaactccaagaac 5869
Db 3126 gattcatgtactgtacaccccttacttatagatggtcgtgcccagtggtggaactccaagaac 3185
QY 5870 aattcaaaaggaagacactcttcgactagctcagatcagccttcccttatattbaaacaagg 5929
Db 3186 aattcaaaaggaagacactcttcgactagctcagatcagccttcccttatattbaaacaagg 3245
QY 5930 tcaatgtcactcaataaagaagaatcatcttaacacaaatgaaagtgttcatltgagagaca 5989
Db 3246 tcaatgtcactcaataaagaagaatcatcttaacacaaatgaaagtgttcatltgagagaca 3305
QY 5990 tgcagaaaaagacacagagatgtgcatcttgcacaacatcagaatcccgagaccaccaaa 6049
Db 3306 tgcagaaaaagacacagagatgtgcatcttgcacaacatcagaatcccgagaccaccaaa 3365
QY 6050 tgcctcagatgttactccagagatctgtfagcacacacagtgaaatcagggtgcttggag 6109
Db 3366 tgcctcagatgttactccagagatctgtfagcacacacagtgaaatcagggtgcttggag 3425
QY 6110 ttgcccaggtttttctgtctgaataactagtgaaaccaaagctcttcagacatcatata 6169
Db 3426 ttgcccaggtttttctgtctgaataactagtgaaaccaaagctcttcagacatcatata 3485
QY 6170 aactcagactcgtccttaagaatcttaactaaaggtgtggaatgaccttaagaanaaata 6229
Db 3486 aactcagactcgtccttaagaatcttaactaaaggtgtggaatgaccttaagaanaaata 3545
QY 6230 agagcttaattgggcgggttcaaaaaggagtccaagaaggatttggggaaatcatcttcgc 6289
Db 3546 agagcttaattgggcgggttcaaaaaggagtccaagaaggatttggggaaatcatcttcgc 3605
QY 6290 cttaagaaggcccttaacaagccctagatccaagaagaagtcctcctaagtatccaagccaagt 6349
Db 3606 cttaagaaggcccttaacaagccctagatccaagaagaagtcctcctaagtatccaagccaagt 3665
QY 6350 gctgtccctcgtccacaagaatctcctcagtcgaaatgagcttgcgcaaaatgagctctcta 6409
Db 3666 gctgtccctcgtccacaagaatctcctcagtcgaaatgagcttgcgcaaaatgagctctcta 3725

QY 6410 aactgaatgcattgttttattcattcctcgaagaagccattatccaacatcgagtga 6469
|||||
Db 3726 aactgaatgcattgttttattcattcctcgaagaagccattatccaacatcgagtga 3785
QY 6470 aagatctatggaacacaacatggaatggaattctggaattattcattcattgaagaatg 6529
|||||
Db 3786 aagatctatggaacacaacatggaatggaattctggaattattcattcattgaagaatg 3845
QY 6530 cagggcgcaagaagaataatcgaatgtaattgtaacgcttggaatcatgctatggtt 6589
|||||
Db 3846 cagggcgcaagaagaataatcgaatgtaattgtaacgcttggaatcatgctatggtt 3905
QY 6590 ctatgttcggttaacaagctgttatctttaaagacattttaaagctcgaagctcact 6649
|||||
Db 3906 ctatgttcggttaacaagctgttatctttaaagacattttaaagctcgaagctcact 3965
QY 6650 ataccattaccattattatcacaatgactaaggttaaaaaattattcattcattgaatcgt 6709
|||||
Db 3966 ataccattaccattattatcacaatgactaaggttaaaaaattattcattcattgaatcgt 4025
QY 6710 atttttaattatataccattattatgattcatttggaccattttaaattgaat 6769
|||||
Db 4026 atttttaattatataccattattatgattcatttggaccattttaaattgaat 4085
QY 6770 gcttattttaaggtactaaaaaatatgtgaatgtttactcctgtgcgcgaagcctc 6828
|||||
Db 4086 gcttattttaaggtactaaaaaatatgtgaatgtttactcctgtgcgcgaagcctc 4144

RESULT 4
AAH41953
ID AAH41953 standard; DNA: 4200 BP.

AC AAH41953;

DT 30-AUG-2001 (first entry)

DE CLASP-3 nucleotide fragment.

Human: cadherin-1-like asymmetry protein; CLASP; CLASP-3; Immune response;
cell surface molecule; transmembrane protein; immunosuppressive; vaccine;
antiinflammatory; antiallergic; antianemic; dermatological; uropathic;
ophthalmological; antirheumatic; nephrotropic; antithyroid; antidiabetic;
neuroprotective; antisthmatic; antibacterial; antisense therapy;

gene therapy: chromosome 1p31.1; ds.

Unidentified.

MO200142297-A2.

14-JUN-2001.

13-DEC-2000; 2000WO-US34171.

13-DEC-1999; 99US-0170453.

14-JAN-2000; 2000US-0176195.

14-FEB-2000; 2000US-0182286.

11-APR-2000; 2000US-0196287.

11-APR-2000; 2000US-0196460.

11-APR-2000; 2000US-0196527.

11-APR-2000; 2000US-0196528.

11-APR-2000; 2000US-0547276.

13-OCT-2000; 2000US-0240503.

13-OCT-2000; 2000US-0240508.

(ARBO-) ARBOR VITA CORP.

Lu P, Garman JD, Candia AF.

WPI: 2001-375003/39.

Novel isolated cadherin-like asymmetry protein (CLASP) -3, useful for

PT useful in preventing or treating a CLASP-3-mediated disease e.g.
PT autoimmune disease
XX
PS Disclosure: Fig 9C; 189pp; English.

CC The present invention describes an isolated polypeptide (I) comprising a
CC nucleotide (nt) sequence (S1a) that has at least 90% identity to the
CC sequence given in AAH41934 and is immunologically cross-reactive with
CC the derived amino acid (aa) sequence (S1b) given in AAB99495 or shares a
CC biological function with native CLASP-3. (I) has immunosuppressive,
CC antiinflammatory, antiallergic, antianemic, dermatological, uropathic,
CC ophthalmological, antirheumatic, nephrotropic, antithyroid, antidiabetic,
CC neuroprotective, antisthmatic and antibacterial activities, and can be
CC used in antisense therapy, vaccine production and gene therapy. CLASP-3
CC related sequences can be used in preventing or treating a CLASP-3
CC mediated disease, preferably an autoimmune disease by inhibiting an
CC immune response. The autoimmune disease is caused or exacerbated by
CC increased activity of T helper 1 lymphocytes (TH1). AAH41893 to AAH41953
CC and AAB99491 to AAB99507 represent sequences which are used in the
CC exemplification of the present invention. CLASP-3 is localised in the
CC chromosome location 1p31.1.

SQ Sequence 4200 BP; 1262 A; 844 C; 907 G; 1187 T; 0 other;

Query Match 60.4%; Score 4123; DB 22; Length 4200;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4136; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 2693 taatcgtctcgaagccttagtaataagcaatccagatatatctggaatcccaagtcac 2752
|||||
Db 61 taatcgtctcgaagccttagtaataagcaatccagatatatctggaatcccaagtcac 120
QY 2753 cagatgataaagttcgatcaatcaccggaaggaaggttttagatcgctcccaattcctg 2812
|||||
Db 121 cagatgataaagttcgatcaatcaccggaaggaaggttttagatcgctcccaattcctg 180
QY 2813 ttaacactgtgtgccaagaagctgcccacatggggaatcccaaccacgaagctgaagat 2872
|||||
Db 181 ttaacactgtgtgccaagaagctgcccacatggggaatcccaaccacgaagctgaagat 240
QY 2873 caacaaggtcattggaatcgaatgttgaatcgtatctctgcacacagaagcgtcaagt 2932
|||||
Db 241 caacaaggtcattggaatcgaatgttgaatcgtatctctgcacacagaagcgtcaagt 300
QY 2933 tcttacaacattcaacggagcgttacaactcaaaaagccttttcaaggaagctgctc 2992
|||||
Db 301 tcttacaacattcaacggagcgttacaactcaaaaagccttttcaaggaagctgctc 360
QY 2993 tgcagtgggtttgttgcgaatgacggttcgggaatcagcttgcacaagcctggtctc 3052
|||||
Db 361 tgcagtgggtttgttgcgaatgacggttcgggaatcagcttgcacaagcctggtctc 420
QY 3053 ttcttgaattatgtaagaagcagtgccacatttacttacttaataacttgaag 3112
|||||
Db 421 ttcttgaattatgtaagaagcagtgccacatttacttacttaataacttgaag 480
QY 3113 ctccaaggaagaagtcgtttccagaagcttcaatggaatgacatgacgtcttgcagca 3172
|||||
Db 481 ctccaaggaagaagtcgtttccagaagcttcaatggaatgacatgacgtcttgcagca 540
QY 3173 cgatgctatgatatagtttcaagatttcagaagaagaagaagaatggttgaagaccca 3232
|||||
Db 541 cgatgctatgatatagtttcaagatttcagaagaagaagaagaatggttgaagaccca 600
QY 3233 atacaagccttgatcttctccaatgacgtcttgcgttatgagcagaagattgctt 3292
|||||
Db 601 atacaagccttgatcttctccaatgacgtcttgcgttatgagcagaagattgctt 660
QY 3293 ttacgcttaaaagtcctgctataaagcagtgcttcaaaagccttaccattaccgaatc 3352
|||||
Db 661 ttacgcttaaaagtcctgctataaagcagtgcttcaaaagccttaccattaccgaatc 720

QY 3353 ccaagtgtctgtgtctcttgagagctgtgatttcttaacgaatcatctgcagtcagtcagtcact 3412
|||||
Db 721 ccaagtgtctgtgtctcttgagagctgtgatttcttaacgaatcatctgcagtcagtcagtcact 780
QY 3413 atgttaccatnaaaccttaacctgcagcttactactccaacctgcattccatcaacctcttg 3472
|||||
Db 781 atgttaacatnaaaccttaacctgcagcttactactccaacctgcattccatcaacctcttg 840
QY 3473 ttctctcttgnaaacatcttcagatgtctgagatttcttaacgaatgtacagaacccaagaattg 3532
|||||
Db 841 ttctctcttgnaaacatcttcagatgtctgagatttcttaacgaatgtacagaacccaagaattg 900
QY 3533 caaataatgttgnaaatatccgtgtccttccgcgaacagcatatatttggaagagactgtgtg 3592
|||||
Db 901 caaataatgttgnaaatatccgtgtccttccgcgaacagcatatatttggaagagactgtgtg 960
QY 3593 taacagagctgtgctgtcaatttttagaacctgtatgtcgaagaagactgttttgatgtacaa 3652
|||||
Db 961 taacagagctgtgctgtcaatttttagaacctgtatgtcgaagaagactgttttgatgtacaa 1020
QY 3653 aagttcatcaatgtgtaacaaatattactctccagtcagatcagacccgcgtactctg 3712
|||||
Db 1021 aagttcatcaatgtgtaacaaatattactctccagtcagatcagacccgcgtactctg 1080
QY 3713 accctcagataaagagctgcaggtgagccagtgtgtatctacacctgtatgttatatcatg 3772
|||||
Db 1081 accctcagataaagagctgcaggtgagccagtgtgtatctacacctgtatgttatatcatg 1140
QY 3773 aaactgtacctcagctgtatgtattttaacagaacatcaacatcaagaagagaaccaattt 3832
|||||
Db 1141 aaactgtacctcagctgtatgtattttaacagaacatcaacatcaagaagagaaccaattt 1200
QY 3833 gttatagccactgtatgtattatgtaaaagtgtgagcggaagtatgataaagccgaacccgtttgca 3892
|||||
Db 1201 gttatagccactgtatgtattatgtaaaagtgtgagcggaagtatgataaagccgaacccgtttgca 1260
QY 3893 tggcaatcgcagagagacatcgtctccctcaactaaacaaagccctggcagtttccctccacgt 3952
|||||
Db 1261 tggcaatcgcagagagacatcgtctccctcaactaaacaaagccctggcagtttccctccacgt 1320
QY 3953 caacgagctgagcagcaacaaactactcttccagcagaatcaagtcgaagccctttgatct 4012
|||||
Db 1321 caacgagctgagcagcaacaaactactcttccagcagaatcaagtcgaagccctttgatct 1380
QY 4013 gtctactgtggttctcnaaaaatgcagatgaacacgttctcaagaaggtgtttacagatc 4072
|||||
Db 1381 gtctactgtggttctcnaaaaatgcagatgaacacgttctcaagaaggtgtttacagatc 1440
QY 4073 tctcagctctgcagcttaaacccggtatatactgtcttactctctgtgtgtctgtctg 4132
|||||
Db 1441 tctcagctctgcagcttaaacccggtatatactgtcttactctctgtgtgtctgtctg 1500
QY 4133 agtataaagggaaaaaagtgcttgaacgaatgaatagcttgaccccttaagaatcaacaa 4192
|||||
Db 1501 agtataaagggaaaaaagtgcttgaacgaatgaatagcttgaccccttaagaatcaacaa 1560
QY 4193 acatgaagagcaaaagcttgaaagagcttatcttcttggaagcataggtgccaagccaagaattg 4252
|||||
Db 1561 acatgaagagcaaaagcttgaaagagcttatcttcttggaagcataggtgccaagccaagaattg 1620
QY 4253 tacggcgaacccgcagagcaacatctcgagagaagcccatctggaagtgtcctttggaagccaag 4312
|||||
Db 1621 tacggcgaacccgcagagcaacatctcgagagaagcccatctggaagtgtcctttggaagccaag 1680
QY 4313 aaaaattggaagtggaagaaagatagtactactgtgcgtcaaaaacacagagaagccttgaca 4372
|||||
Db 1681 aaaaattggaagtggaagaaagatagtactactgtgcgtcaaaaacacagagaagccttgaca 1740
QY 4373 aatcaagagcagagattggaacacgagacactgtatgtatgnaaaaacccgtgtacagagcaaa 4432
|||||
Db 1741 aatcaagagcagagattggaacacgagacactgtatgtatgnaaaaacccgtgtacagagcaaa 1800
QY 4433 acctaatcatatttagatatacttagagatgtgtctcagaccgttctgttaacgaatca 4492

Db 1801 acctaatcatatttagatatacttagagatgtgtctcagaccgttctgttaacgaatca 1860
QY 4493 aagagaacatctctgtgtgaagtgtctaaagtgtactacacagatagccgtgtacaaca 4552
|||||
Db 1861 aagagaacatctctgtgtgaagtgtctaaagtgtactacacagatagccgtgtacaaca 1920
QY 4553 gtgcagttatctataaacactgtttgtgtacagagagccttggtttcaagttctgt 4612
|||||
Db 1921 gtgcagttatctataaacactgtttgtgtacagagagccttggtttcaagttctgt 1980
QY 4613 aactctatttgaagaagaagaagcagtgctgtcatltaigtccoaagcttccagac 4672
|||||
Db 1981 aactctatttgaagaagaagaagcagtgctgtcatltaigtccoaagcttccagac 2040
QY 4673 actgtagcagtagcaatcgtgtacaataaggtcaacacccagtgctcccttaactactaa 4732
|||||
Db 2041 actgtagcagtagcaatcgtgtacaataaggtcaacacccagtgctcccttaactactaa 2100
QY 4733 tgaagcaaaacttgaagattggaatcttgccaaggttaaaaatgcagttacaatgt 4792
|||||
Db 2101 tgaagcaaaacttgaagattggaatcttgccaaggttaaaaatgcagttacaatgt 2160
QY 4793 cactatccctctgtgtgagcacaatccagaaatttaatgaagaattcttaagacgttctc 4852
|||||
Db 2161 cactatccctctgtgtgagcacaatccagaaatttaatgaagaattcttaagacgttctc 2220
QY 4853 taaagaactatattgaacataatgtcgaagaagactgtgaagtgaagggaacaacattctgt 4912
|||||
Db 2221 taaagaactatattgaacataatgtcgaagaagactgtgaagtgaagggaacaacattctgt 2280
QY 4913 atcaggtccagagatctgttttccaatccatcgaatgtcttcctgtacactgtgaatga 4972
|||||
Db 2281 atcaggtccagagatctgttttccaatccatcgaatgtcttcctgtacactgtgaatga 2340
QY 4973 aggaacacccaagagagatctctgaatgtgtatgtactaagttaacagaatgtccaaggtt 5032
|||||
Db 2341 aggaacacccaagagagatctctgaatgtgtatgtactaagttaacagaatgtccaaggtt 2400
QY 5033 accaagactctccaga----gagattgacccgtgtgcagaacaatgcaggaagcactccag 5089
|||||
Db 2401 accaagactctccagaatctgcagatgtacctgtgtgcagaacaatgcaggaagcactccag 2460
QY 5090 aacgaagcaatcatgtcgaagctgtcagacagtgctcagaccctccagcaacattgtctg 5149
|||||
Db 2461 aacgaagcaatcatgtcgaagctgtcagacagtgctcagaccctccagcaacattgtctg 2520
QY 5150 aatattgagcagtctggaagaccggaataatctctcctgtggaatgttlaacattcaga 5209
|||||
Db 2521 aatattgagcagtctggaagaccggaataatctctcctgtggaatgttlaacattcaga 2580
QY 5210 atatttcatctaaatgttttaagaagaatctgcggtctcagatgatgtgtatctccagatg 5269
|||||
Db 2581 atatttcatctaaatgttttaagaagaatctgcggtctcagatgatgtgtatctccagatg 2640
QY 5270 aagaaggtatctgtccttgaaatacttactgaatccagagcctgttgagatctacatgaac 5329
|||||
Db 2641 aagaaggtatctgtccttgaaatacttactgaatccagagcctgttgagatctacatgaac 2700
QY 5330 aagcagctgtctcctctctatagctgtgcagatgaatgaagcagttlaatgaattlacaaag 5389
|||||
Db 2701 aagcagctgtctcctctctatagctgtgcagatgaatgaagcagttlaatgaattlacaaag 2760
QY 5390 tacttatctctatcaatgaagcgaatcgtgagatgaagaagaacatccaaatcaatgtga 5449
|||||
Db 2761 tacttatctctatcaatgaagcgaatcgtgagatgaagaagaacatccaaatcaatgtga 2820
QY 5450 aacttcaagaagcattcagcaaaatgttcatccaagaagcagcgtcggagagcgagattgt 5509
|||||
Db 2821 aacttcaagaagcattcagcaaaatgttcatccaagaagcagcgtcggagagcgagattgt 2880
QY 5510 gacattattctgtgtgttttlaatgaacaaagtctcgagatttgatgaacaaat 5569
|||||

Dh 2881 gcacctatttcgtgtgtgttttatttgaaccaagttcggggatttgatgaacaagaat 2940
Qy 5570 ttgtttacaagagagccttgcacaataactttgcagagatatctcacagattgaggat 5529
Dh 2941 ttgtttacaagagagccttgcacaataactttgcagagatatctcacagattgaggat 3000
Qy 5630 ttacgagaaagaatttggagagatgttggttgaagtaacaagaagacttcaatccttga 5689
Dh 3001 ttacgagaaagaatttggagagatgttggttgaagtaacaagaagacttcaatccttga 3060
Qy 5690 acaagtgtaaatatgatctcacaagacatatctcagattactcctatgttgagccatcact 5749
Dh 3061 acaagtgtaaatatgatctcacaagacatatctcagattactcctatgttgagccatcact 3120
Qy 5750 ttgcacataatgatgaagagacagaatcaacctttccgcaaaaattacatcttcgtc 5809
Dh 3121 ttgcacataatgatgaagagacagaatcaacctttccgcaaaaattacatcttcgtc 3180
Qy 5810 gattcaatgtactgtacaccccttactttagatgacg tgcacatgaggaaacttcaatgaac 5869
Dh 3181 gattcaatgtactgtacaccccttactttagatgacg tgcacatgaggaaacttcaatgaac 3240
Qy 5870 aattcaaaaggaagacacattctgaactcagtcacatgccttcttataataaacaagg 5929
Dh 3241 aattcaaaaggaagacacattctgaactcagtcacatgccttcttataataaacaagg 3300
Qy 5920 tcaatgttactcataaagaagaatcacttaacacaaatttgatgttcattgagagaca 5989
Dh 3301 tcaatgttactcataaagaagaatcacttaacacaaatttgatgttcattgagagaca 3360
Qy 5990 tgcagaaaaagacacagagatgttcgacatttgcacacatcagatcccgagaccccaaaa 6049
Dh 3361 tgcagaaaaagacacagagatgttcgacatttgcacacatcagatcccgagaccccaaaa 3420
Qy 6050 tgcctcagaatgtatctcacaagagatctgtaggcaccacagtgaaatcagggtccttggag 6109
Dh 3421 tgcctcagaatgtatctcacaagagatctgtaggcaccacagtgaaatcagggtccttggag 3480
Qy 6110 ttgcccaggttttctgtctgaataatctagtgagcccaagcctttagacatcatata 6169
Dh 3401 ttgcccaggttttctgtctgaataatctagtgagcccaagcctttagacatcatata 3340
Qy 6170 aactcgcagctcgtcttaagaatttactaaagggtgtgaagatgaccttaagaanaata 6229
Dh 3541 aactcgcagctcgtcttaagaatttactaaagggtgtgaagatgaccttaagaanaata 3600
Qy 6230 agagcttaattggcgcggttcaaaaagatcatcaaaagggaattggggaaactatcttcgc 6289
Dh 3601 agagcttaattggcgcggttcaaaaagatcatcaaaagggaattggggaaactatcttcgc 3660
Qy 6290 cttaaagagggccctacagcccttagatcacagaaagtcctccagttatccaagccagttat 6349
Dh 3661 cttaaagagggccctacagcccttagatcacagaaagtcctccagttatccaagccagttat 3720
Qy 6350 gctgttcccttgcacagagatttccctcagtcgaatgagcttctgcanaaattgagatctcta 6409
Dh 3721 gctgttcccttgcacagagatttccctcagtcgaatgagcttctgcanaaattgagatctcta 3780
Qy 6410 aacggaatgcactgttttattatctctgcanaagagccatgtatccaacatcgatgtgaa 6469
Dh 3781 aacggaatgcactgttttattatctctgcanaagagccatgtatccaacatcgatgtgaa 3840
Qy 6470 aagatctatggaaccaacaatggaatggaattctgnaattattattatcttgaagaatg 6529
Dh 3841 aagatctatggaaccaacaatggaatggaattctgnaattattattatcttgaagaatg 3900
Qy 6530 caatgtgcgaagaataataatcaaatgtagatgttgaacgcttggaatcatgtctatggttt 6589
Dh 3901 caatgtgcgaagaataataatcaaatgtagatgttgaacgcttggaatcatgtctatggttt 3960
Qy 6590 ctaatgttcgggttaacaagcgttatcttttaagaacatttaatgactcaaaagttacact 6649
Dh 3961 ctaatgttcgggttaacaagcgttatcttttaagaacatttaatgactcaaaagttacact 4020

Qy 6650 atacatttaccattattatcacatagcgaagggttaaaaattattcactttaagttcgt 6709
Dh 4021 atacatttaccattattatcacatagcgaagggttaaaaattattcactttaagttcgt 4080
Qy 6710 atttttaattatatacacatttataagattcattttggaccattttaaatgtagtaat 6769
Dh 4081 atttttaattatatacacatttataagattcattttggaccattttaaatgtagtaat 4140
Qy 6770 gctatttttaagggtactaaaaatagtgaatgtttactctgtgcggcgagggccctc 6828
Dh 4141 gctatttttaagggtactaaaaatagtgaatgtttactctgtgcggcgagggccctc 4199

RESULT 5
AAH41952
ID AAH41952 standard; DNA; 3173 BP.
XX
AC AAH41952;
XX
DT 30-AUG-2001 (first entry)
XX
DE CLASP-3 nucleotide fragment.
XX
XX Human; cadherin-like asymmetry protein; CLASP; CLASP-3; immune response;
KW cell surface molecule; transmembrane protein; immunosuppressive; vaccine;
KW antiinflammatory; antiarthritic; antianemic; dermatological; uropathic;
KW ophthalmological; antitubercular; nephrotoxic; antidiabetic;
KW neuroprotective; antiaslatic; antibacterial; antisense therapy;
KW gene therapy; chromosome 1p31.1; ds.
XX
OS Unidentified.
XX
XX WO200142297-A2.
PN 14-JUN-2001.
XX
PD 13-DEC-2000; 2000WO-US34171.
XX
PF 13-DEC-1999; 98US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
PR 11-APR-2000; 2000US-0196267.
PR 11-APR-2000; 2000US-0196460.
PR 11-APR-2000; 2000US-0196527.
PR 11-APR-2000; 2000US-0196528.
PR 11-APR-2000; 2000US-0547276.
PR 13-OCT-2000; 2000US-0240503.
PR 13-OCT-2000; 2000US-0240508.
XX
XX (ARBO-) ARBOR VIRA CORP.
PA
XX
XX Lu P, Garman JD, Candia AF;
PI
XX
XX WPI; 2001-375003/39.
DR
XX
XX Novel isolated cadherin-like asymmetry protein (CLASP) -3, useful for
PT useful in preventing or treating a CLASP-3-mediated disease e.g.
PT autoimmune disease
PS
XX
XX Disclosure; Fig 9A; 189pp; English.
XX
XX The present invention describes an isolated polypeptide (I) comprising a
XX nucleotide (nt) sequence (S1a) that has at least 90% identity to the
XX sequence given in AAH41934 and is immunologically cross-reactive with
XX the derived amino acid (aa) sequence (S1b) given in AAB99495 or shares a
XX biological function with native CLASP-3. (I) has immunosuppressive,
XX antiinflammatory, antiarthritic, antianemic, dermatological, uropathic,
XX ophthalmological, antitubercular, nephrotoxic, antidiabetic,
XX neuroprotective, antiaslatic and antibacterial activities, and can be
XX used in antisense therapy, vaccine production and gene therapy. CLASP-3
XX related sequences can be used in preventing or treating a CLASP-3
XX mediated disease, preferably an autoimmune disease by inhibiting an

CC Immune response. The autoimmune disease is caused or exacerbated by
CC increased activity of T helper 1 lymphocytes (TH1). AHH1893 to AHH1953
CC and ABB9491 to ABB9507 represent sequences which are used in the
CC exemplification of the present invention. CLASP-3 is localised in the
CC chromosome location 1p31.1.

XX Sequence 3173 BP; 981 A; 624 C; 688 G; 880 T; 0 other;

Query Match 44.6%; Score 3044.8; DB 22; Length 3173;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 3155; Conservative 0; Mismatches 12; Indels 83; Gaps 2;

QY 3582 aggaatttggttaacaagagcgcgcgtgcatattttagacctgagtctgaagagctgttg 3641
DB 3 aggaactgtgttttcagagcgtgcgtgcatattttagacctgagtctgaagagctgttg 62
QY 3642 attgataagaagatcatalatggtacacaatttactctcagtcacagatcagacc 3701
DB 63 attgataagaagatcatalatggtacacaatttactctcagtcacagatcagacc 122
QY 3702 gcggtactctgacctgaataagcgtcagctgagtcgcatgtgtatctactctgtattg 3761
DB 123 gcggtactctgacctgaataagcgtcagctgagtcgcatgtgtatctactctgtattg 182
QY 3762 tatatcatggaactgctgactcagctgtatgtattttaaagaaactcaaatcaagcag 3821
DB 183 tatatcatggaactgctgactcagctgtatgtattttaaagaaactcaaatcaagcag 242
QY 3822 aagaccaatttggtatagccactgatatatgaaagtgaagcgggaagtatgataagcca 3881
DB 243 aagaccaatttggtatagccactgatatatgaaagtgaagcgggaagtatgataagcca 302
QY 3882 gaccgttaccatggaatcgcagagacatcgctccctcaataaagaagcctgscattt 3941
DB 303 gaccgttaccatggaatcgcagagacatcgctccctcaataaagaagcctgscattt 362
QY 3942 cctctcaccgtcaacagatgagcagcaaacacataccttttcaagaatcaaatctgaag 4001
DB 363 cctctcaccgtcaacagatgagcagcaaacacataccttttcaagaatcaaatctgaag 422
QY 4002 cctttgtctgctcactcttggtgttctcaaaatgacagatgaacaagttctcaagaagt 4061
DB 423 cctttgtctgctcactcttggtgttctcaaaatgacagatgaacaagttctcaagaagt 482
QY 4062 gtttaacaatctcagcttcgacgtctaaacggcctattagatctgtcttactctgtgc 4121
DB 483 gtttaacaatctcagcttcgacgtctaaacggcctattagatctgtcttactctgtgc 542
QY 4122 gtctgtcttgagataaagaggaaaaagtgtttgaacgaatgaatagcttgaccctttaa 4181
DB 543 gtctgtcttgagataaagaggaaaaagtgtttgaacgaatgaatagcttgaccctttaa 602
QY 4182 gaaatcaaaaacatgagagcaaaagcttggaagaagctattcttggagcagatagtcgcag 4241
DB 603 gaaatcaaaaacatgagagcaaaagcttggaagaagctattcttggagcagatagtcgcag 662
QY 4242 gcaagaatattgtaagcgcgaagccgagacagctcgagagaagcccatctggaagtgcctt 4301
DB 663 gcaagaatattgtaagcgcgaagccgagacagctcgagagaagcccatctggaagtgcctt 722
QY 4302 tggaaagtcaagaanaatttgaggctggaagaaagatagaactcactggtgcgtcaaaacacaga 4361
DB 723 tggaaagtcaagaanaatttgaggctggaagaaagatagaactcactggtgcgtcaaaacacaga 782
QY 4362 gaagcttgaacaatcaagaagcagagatgtgaacacgaagacactgattatggaacactggc 4421
DB 783 gaagcttgaacaatcaagaagcagagatgtgaacacgaagacactgattatggaacactggc 842
QY 4422 tacagaagaacaacctaatcatttagatatactagagattgttctgaagcgtttctgt 4481
DB 843 tacagaagaacaagtaatcatttagatatactagagattgttctgaagcgtttctgt 902

QY 4482 aacgaatccaaagagagcatcttggttgagtgctaaagtgtactacacagcatgac 4541
DB 903 aacgaatccaaagagagcatcttggttgagtgctaaagtgtactacacagcatgac 962
QY 4542 ctgttaaccaagtgtcagttatctacaaacgttttctgaacagagagccttggttc 4601
DB 963 ctgttaaccaagtgtcagttatctacaaacgttttctgaacagagagccttggttc 1022
QY 4602 aaagtttctcgaactcttatttgaagaagaagacagacgttgtctgattatgcctag 4661
DB 1023 aaagtttctcgaactcttatttgaagaagaagacagacgttgtctgattatgcctag 1082
QY 4662 gctctccagacatgtgacgttagcatcggtaacaatggtcaacccagtgctccct 4721
DB 1083 gctctccagacatgtgacgttagcatcggtaacaatggtcaacccagtgctccct 1142
QY 4722 ttacctaactaagagcaaaaacttggagatgggaaataacttgcagaggttaaatgca 4781
DB 1143 ttacctaactaagagcaaaaacttggagatgggaaataacttgcagaggttaaatgca 1202
QY 4782 ggtacaagtgtcactctcctccttggtgggcacatctcgaatttaatgaagaattctt 4841
DB 1203 ggtacaagtgtcactctcctccttggtgggcacatctcgaatttaatgaagaattctt 1262
QY 4842 aagaagtttctctaaagactatattgacatactgtgaagaagactctgaaatgaagaaac 4901
DB 1263 aagaagtttctctaaagactatattgacatactgtgaagaagactctgaaatgaagaaac 1322
QY 4902 aacatttctgtatcaggttccagaatctggttttcaatctccataatgattcttcgtatc 4961
DB 1323 aacatttctgtatcaggttccagaatctggttttcaatctccataatgattcttcgtatc 1382
QY 4962 tgtgaaatgaagaacacacagagagatctctgaatggttgattgattcttaagaat 5021
DB 1383 tgtgaaatgaagaacacacagagagatctctgaatggttgattgattcttaagaat 1442
QY 5022 tgcacaagggttaccagaacactccacaga---gcgattgacctggttgacagaatactgcag 5078
DB 1443 tgcacaagggttaccagaacactccacagaatctgcatgacctggttgacagaatactgcag 1502
QY 5078 caagcactcagaagacgaatcatctgaagctgcacagttctcagctcactcagcagc 5138
DB 1503 caagcactcagaagacgaatcatctgaagctgcacagttctcagctcactcagcagc 1562
QY 5139 actgtgtgcgaatatttgaacgtctgtagagacgggaatatctccctggtggagtgt 5198
DB 1563 actgtgtgcgaatatttgaacgtctgtagagacgggaatatctccctggtggagtgt 1622
QY 5199 aacatttcagaatatttcaactaattttagaagaatctgcgttcagatgattgtgt 5258
DB 1623 aacatttcagaatatttcaactaattttagaagaatctgcgttcagatgattgtgt 1682
QY 5259 atctccaagatgaagaaggtatctgctctcgtgaaaaatacttactgagtcagaagctgtgg 5318
DB 1683 atctccaagatgaagaaggtatctgctctcgtgaaaaatacttactgagtcagaagctgtgg 1742
QY 5319 attacgtggaacaagcagcgtctcctcctcctatggtgcagtgaatgaagcagtttaata 5378
DB 1743 attacgtggaacaagcagcgtctcctcctcctatggtgcagtgaatgaagcagtttaata 1802
QY 5379 agtttaacaagtaactatctcctatltcaatgaagctaaatctgggaatgcgaanaactatccac 5438
DB 1803 agtttaacaagtaactatctcctatltcaatgaagctaaatctgggaatgcgaanaactatccac 1862
QY 5439 aattcaatgtgaaactccaagaagcatctcagcaaaattgttcaatcagagtaactgctgtgga 5498
DB 1863 aattcaatgtgaaactccaagaagcatctcagcaaaattgttcaatcagagtaactgctgtgga 1922
QY 5499 gcggatgttggacactatttcgtgtgtttttagaacaacaagtttgggatttga 5558
DB 1923 gcggatgttggacactatttcgtgtgtttttagaacaacaagtttgggatttga 1982
QY 5559 tgaacaagaattgttttaagaaggcctgtcaataaacaacttgcagagatactcacag 5618

|||||
Db 1983 tgaacaagaattggtttacaagagcctgcataaaccnaacttgcagagatattccaag 2042
QY 5619 attggaaggattttacggaagaatttggagaga tgggttgaagtaatccaagactc 5678
Db 2043 attggaaggattttacggaagaatttggagagtgggttgaagtaatccaagactc 2102
QY 5679 taatcctgaaagaacagtgttaattagatccctaacaagacatatcttcagattactagt 5738
Db 2103 taatcctgaaagaacagtgttaattagatccctaacaagacatatcttcagattactagt 2162
QY 5739 ggaagcactacttgcacacatgagatgaaagacagaaatcactatttcgacaaaatla 5798
Db 2163 ggaagcactacttgcacacatgagatgaaagacagaaatcactatttcgacaaaatla 2222
QY 5799 caattctgcgtacatcgtactgtacaccccttactttagatggcgtgccatgggga 5858
Db 2223 caattctgcgtacatcgtactgtacaccccttactttagatggcgtgccatgggga 2282
QY 5859 acttatgaaacaattcaaaagaagacacattctgactcgtctcactgcttccctatat 5918
Db 2283 acttatgaaacaattcaaaagaagacacattctgactcgtctcactgcttccctatat 2342
QY 5919 taaaacaaggtgcaatgtcactcacaagaagaatcacttaacaccaattgaagtgc 5978
Db 2343 taaaacaaggtgcaatgtcactcacaagaagaatcacttaacaccaattgaagtgc 2402
QY 5979 tattgaagacatgacgaagaagaacacagagatttgcatttgcacacatcagatccgc 6038
Db 2403 tattgaagacatgacgaagaagaacacagagatttgcatttgcacacatcagatccgc 2462
QY 6039 agaccaccaaatgtctcagatgtactcagagatctgtaagcacacagatgaaatcag 6098
Db 2463 agaccaccaaatgtctcagatgtactcagagatctgtaagcacacagatgaaatcag 2522
QY 6099 gcccttggaaagtgtcccaaggtttctgttctgnaaatacctagtgacccaagctctcag 6158
Db 2523 gcccttggaaagtgtcccaaggtttctgttctgnaaatacctagtgacccaagctctcag 2582
QY 6159 acatatataaactgcgactctgtcttaagaatttacttaaaagtggtgaagatgctt 6218
Db 2583 acatatataaactgcgactctgtcttaagaatttacttaaaagtggtgaagatgctt 2642
QY 6219 aagaaaaataagagcttaatttggccggttcaaaagagatcacaagggatctgggaa 6278
Db 2643 aagaaaaataagagcttaatttggccggttcaaaagagatcacaagggatctgggaa 2702
QY 6279 actatcttcgccttaagagagcctacagccctagatcacagaaagtcctcagtattcc 6338
Db 2703 actatcttcgccttaagagagcctacagccctagatcacagaaagtcctcagtattcc 2762
QY 6339 aagcagatattgtctgtccctgcacagagatcttccctcagtcgaatgagcttccgcaa 6398
Db 2763 aagcagatattgtctgtccctgcacagagatcttccctcagtcgaatgagcttccgcaa 2822
QY 6399 atgatactcctaactgaatgcactgttttattcatctgcacaaagggccatctatcaaca 6458
Db 2823 atgatactcctaactgaatgcactgttttattcatctgcacaaagggccatctatcaaca 2882
QY 6459 tcgagtcgtgaaaaagatctatttggaaaccaacatggaatcgtgaaattattatttc 6518
Db 2883 tcgagtcgtgaaaaagatctatttggaaaccaacatggaatcgtgaaattattatttc 2942
QY 6519 attgaagaatgtagtgcccaagaataatcaatgttagatgttaacgcttgagaatcat 6578
Db 2943 attgaagaatgtagtgcccaagaataatcaatgttagatgttaacgcttgagaatcat 2959
QY 6579 ggctatggtttctaattgttcggtatacaagcgttattctttaagacatttaagtactc 6638
Db 2960 -----ctttaagacattttaagtactc 2962
QY 6639 aaaggtacatatatattaccattattatataccaatgaagttaaaaattattacc 6698
|||||

Db 2983 aaaggtacactatagatttaccattattatataccatagctaaagttaaaattattacc 3042
QY 6699 ttttaagtcgtatttttattatataccattatataatcatttatttggaccatttta 6758
Db 3043 ttttaagtcgtatttttattatataccattatataatcatttatttggaccatttta 3102
QY 6759 aatgtagtaatgcttattttaagaagtaactaaaataatgtaagtttaccctggtcgcg 6818
Db 3103 aatgtagtaatgcttattttaagaagtaactaaaataatgtaagtttaccctggtcgcg 3162
QY 6819 ccagggcctc 6828
Db 3163 ccagggcctc 3172

RESULT 6
AAH43851
ID AAH43851 standard; cDNA; 6372 BP.
XX
AC AAH43851:
XX
DT 04-SEP-2001 (first entry)
XX
DE Human CLASP-7 encoding cDNA sequence SEQ ID NO:1.
XX
KW Human: CLASP-7; cadherin-like asymmetry protein; immune response;
KW neuroprotective; antidiabetic; immunosuppressive; antirheumatic;
KW antitubercle; hypotensive; anti-HIV; cytostatic; immunostimulant;
KW antianemic; antinflammatory; ophthalmological; nephrotoxic;
KW antihydroid; antiallergic; antibacterial; gene therapy;
KW chromosome 19q13.2; autoimmune disease; multiple sclerosis; toxemia;
KW juvenile diabetes; rheumatoid arthritis; pruritic urticarial papule;
KW hypertension; Rh incompatibility; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 13..6156
FT /tag= a
FT /product= "CLASP-7"
FT
XX
PN WO200142295-A2.
XX
PD 14-JUN-2001.
XX
PE 13-DEC-2000; 2000MC-US34152.
XX
PR 13-DEC-1999; 990S-0170453.
XX
PR 14-JAN-2000; 2000US-0176195.
XX
PR 14-FEB-2000; 2000US-0182296.
XX
PR 11-APR-2000; 2000US-0196267.
XX
PR 11-APR-2000; 2000US-0196460.
XX
PR 11-APR-2000; 2000US-0196527.
XX
PR 11-APR-2000; 2000US-0196528.
XX
PR 11-APR-2000; 2000US-0547276.
XX
PR 13-OCT-2000; 2000US-0240503.
XX
PR 13-OCT-2000; 2000US-0240508.
XX
PR 13-OCT-2000; 2000US-0240539.
XX
PR 13-OCT-2000; 2000US-0240543.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu P, Garman JD, Candia AF;
XX
DR WPI; 2001-381641/40.
XX
DR P-PsDB; AAB9541.
XX
XX
XX Novel cadherin-like asymmetry protein-7 and polynucleotides encoding
XX the polypeptide, useful for treating autoimmune disease,
XX hypersensitivity, preventing transplant rejection by modulating immune
XX response -
XX
PS Claim 1; Fig 5; 151pp; English.

QY	1766	ataataacagtgaagaagccaggttatatgatatgagagagatcccaagcaatgcatacccggttaa	1825
QY	1767	atataacagctgaagaagccaggttatatgatatgagagagatcccaagcaatgcatacccggttaa	1826
Db	1718	acctctgtctgtgcagatgtgcagttacatgatcacagcgccagagaccccgacagcgctctgcggttca	1777
QY	1826	tctctgttgaactctagctgtgtcttgcaaatcttccaagaagccataccagcgtaatatc	1885
Db	1778	tctctgttgaactctagctgtgtcttgcaaatcttccaagaagccataccagcgtaatatc	1887
QY	1886	ataacagctctcctgatatttcatalgaaagaatcacaaggttaagcttcctgctaacttaactg	1945
Db	1838	ataacagaagctcccgaggttctacagagagattcaagctgactcttcacagcctgcgtgcagag	1897
QY	1946	accatactcaactgtgcttttaacttttatacatggttagttgttcaaaaaaataaataactc	2005
Db	1898	agaaacccctacacgtgcgtttcaaccttctacacagttcagctgcagaccccgcccggtgcactg	1957
QY	2006	ctctctgaacacacagatctgtgatatacatatgataccaatgcttcacagatatgacggtttgaaga	2065
Db	1958	ccctgtgagacacacgctgtgaggttactttgtatcccaactgctgtgagaacagcgccgtcgga	2017
QY	2066	ctgtgcacgatttgcgttcgacagcttcattctgaaaaaacacacacagagcttatctgtactgt	2125
Db	2018	ccggcccccctctgtctcccccagtgctctgtgtgagacacgcgcgcgcgcgtacttcgtgtca	2077
QY	2126	ctctctgaggttcccttcaacccgagatgaagaagtgtgtgatatacatacaagaagtgttttaagt	2185
Db	2078	caaccgcatgtgagcgcttcccgagcatatgcgtctgtggttgagacggttcaacaagcggtgttcagt	2137
QY	2186	tgtgaagtgtgtgcgtttcgtctatccatcacacacaagaatcctatctatctgtacaaatttttg	2245
Db	2138	tgtgagctccacaagccggtgtctctctgtgtgcaccccacagacccctacactgtgacaattcttca	2197
QY	2246	ctctctgtccaaatgcctctgtgatagaacacactgttccacagctccgaattgtggagacatgtgaaaca	2305
Db	2198	ccctgtgtgcacgctccctgtgagagagggagcccttcacatcccggtcctcaagaacactgtgtga	2257
QY	2306	tgtgaaaaataactttagaanaatgtgaatgtgaagagagatatttcagacactgtgaatcatccacgc	2365
Db	2258	ggcagggcacaacgtgtgagcagagagctgcgagccagctcttgcaagcactgcgcttcggccagcc	2317
QY	2356	tgtgaacacaggtgtccgaattcttctcatccttcgtctgataataactgtacttttaattat	2425
Db	2318	ccgaacacccctgtgtgccttctccacacacagctgtgagacacgctgtgtgcctgtgtgtcatca	2377
QY	2426	gaacctctgtcatatgtcgtgcacaaatagttaacctgataacgacactctttagaacatgtg	2485
Db	2378	ggccccgcatcatcagatgtgagccagattgtgaacctgtgcgcgtgtgagcctttagaagcaatg	2437
QY	2486	catcaattataaatcgacttcacaaaaactgtgaagagaatcatagtaccagcatgtgcagaa	2545
Db	2438	cccaatgtagatcaactgtgttcaacccgagacgctgtgagagccagcccaagatgtcccgcgctact	2497
QY	2546	acagacctcttgcatactatatacatatcatatgattttccgcgccttaccgaatacttaactaat	2605
Db	2498	ggcccaacgctgtgctgctactaagctccactaagccttccgccttccttcgtgcactgtagccca---	2554
QY	2606	catcatcaccaaggtctcctgtgaggtttgttgagagatcagttgatattgtccacaatgtgtagat	2665
Db	2555	-----gctccccgagatgtggggcccccctccagtgacagtgagcgtgcacacatgtgcgcgtg	2608
QY	2666	ctgtggtgtgagacctgtgcaagccttaatttaaalccggttcttcgtgaagccttgtaatatgcatac	2725
Db	2609	gctctgtgtccgcccgcaagcctctctaacctgtgcggttcccaagagcatcagcagcagcaacc	2688
QY	2726	cagatatctctgtgagatccccaagctccacacagtgatataagattgtgatcaatcatcctggagta	2785
Db	2669	ctgacccctgcggtgtgcccccggtcgtctgtgtgagacagagatttcccgcatcctcgtgcag--	2726
QY	2786	agggtttagatcgtccaattcctctgtgttaaacactgtgtgtccaaaagctgtcccatgtgtg	2845
Db	2727	-----	2746

QY	2846	gatlccaaaccccgatccaaatgtgcagaatccaacaggtctatgtaacgttgtaatcgtta	2905
Db	2727	-----	2726
QY	2906	tgctctgcacacagagagcgtcaagtctcttaacaacattacagggacgttaccacta	2965
Db	2727	-----	2726
QY	2966	aaaagcttcttcacagagagagcttgctcttgcaatgaggtttgttcgacgtgacggttcggg	3025
Db	2727	-caagctgcttcacagagagagcttgctcttgcaatgaggtttgttcgacgtgacggttcggg	2785
QY	3026	aatacagcttgcaacaacagccttgctctcttttggaaattatgtgttaagagcgtgtgcaac	3085
Db	2786	agggcattccctcagcaacagccttgctctctcttcacagctcagtggtgaagatgatgtgcgttcgc	2845
QY	3086	atttatacttaatgataaactctgagggctcccaaggaaagtcggttttcacgaagaagtttca	3145
Db	2846	aactgctgcttggtgcacagcagctcagatacacaccccgcaagctggtcttccccgagacgtttcc	2905
QY	3146	tgaatgacatgtagcgtcctctgtacagacagattgtctatgatatatttccacagatttcaaga	3205
Db	2906	tggagacacatcagctcgttggttggtgctctgtgggctcgtagaggtatcaaccgggtgccaaca	2965
QY	3206	agagcaacaagaatgtgtgagagacatcaatacaagcctgtgcatctcttcctcaatgactcgt	3265
Db	2966	agagatgtgagagcttggtgcagacatccatacagccaagcctggtctcttcctccagtgacttcc	3025
QY	3266	tgctctgttatggagagagaggtttgttttgaagctcttaataagtcctgcgtatataacaggtgt	3325
Db	3026	tgctccctggttgagacgggtgctctgtatccctcaagccttggtccgggcccacatacaagcaggtg	3085
QY	3326	cttcaaaagcttatactcaatlcacagacatccagactgcttggtgtgctcttgtagggcttggaatttc	3385
Db	3086	ccaagcggtctccagatcgtctccccaatccagacagcagcctggttgagcccttgtagcatgtgaattca	3145
QY	3386	tacgaatcatctgcagatcagatgagcactatgtttaacataacttaccctgcagcttaactta	3445
Db	3146	cccgatctcgttgtagcaacagagacactatgtagccctccacatccctccctgctgcctccgt	3205
QY	3446	ctccacactgatactccatcaactcctctgctttctcttgcaacactctcaagagttctctgatttt	3505
Db	3206	cacctccagcctctgcctccctccctccctctgtgtctccccaacaccccccagagctccacactctc	3265
QY	3506	ctacgaaatgtacaagagcaaaagaattgtgcaaatatgttttaattatccgtgcgtcttccgccc	3565
Db	3266	ccagccaaagccccgagaccccacaaagtgtgacacagatgcttgaaacttgagtgagaccattccgcg	3325
QY	3566	aacagcattatttggcagagactgtgtgttaacagagcctggtgtcatcttttaagacctgtag	3625
Db	3326	agcagcaacttccctagctgtggtctctctgtgacggaggtggtgcattggtccctcgaaacctggg	3385
QY	3626	ctgaaagacgtgttttgattttgcaataaagaatcatcaataatgttlaacaacattactctcca	3685
Db	3386	ctgaaaggggcatcttctgtgtgcaaaagaggtcactcagctgtgtgtagcagcgtcgtcatatgtg	3445
QY	3686	gtacagacttgaagaccgggttactctgcagacccctcaagaataagagctgaaatgtagcatatgtt	3745
Db	3446	gcatatgaacatgcagcccccgcttaagccgagagccacgtgtgaaggtcgtgtgtgcccgaagctgt	3505
QY	3746	atctacactgtatgtgtatatacatatgaaagtgtaacctcagctgtatgattttacagaaa	3805
Db	3506	aactgtccactgtcttctgattgtgcaaggtatcacttgcacagctgtgcatgtactttgtctgag	3565
QY	3806	ctcaacaatcaacgaggaagac-----caattgttatagccactgtatgtatgtaaagt	3865
Db	3566	gcccagaggttcacgggttcaagagactgtgctccaatagtcttgaactcaacacagaaagcgaagggg	3625
QY	3860	agagcgggaagatgtatgaataagccagacggttgtagcatgtgcaatctgcaagggacatcggttccctc	3919
Db	3626	acattgtcggtgataccaaccaacccctctgtggccaatgtgccaatgtgcgttgggtgggtcccccagacc	3685
QY	3920	aacttaacaaggtctgtagactttctctcctcaacgttcaacaggtgtgcagggcaacacatacct	3979

Db 425 ccgtgtcagcgagcccccgcacctaacgctgtctgtcgacgtctctgaggaaagcc 484
QY gctggccctgtagatctttagcttgaaataatcacttcctgtacgtcttgcctccaatc 649
Db 485 ccgtacacgtcctgtgactcttgactccgcagccctgcagcctcgaaacggtagaaacc 544
QY tactgtacgaactccaatgaaataagacgcgtcagaatgtatgcacaaagaaatcaa 709
Db 545 tctctcagcaagtgtgtccgagagactttagaagcagaacgagagagcccggtgagaca 604
QY aacgtcacaaagaacttcttctgtcatccatcagatgagaaagaaccaaataaac 769
Db 605 aacagcagccgacgtccttgccttctacccaatgagagagagatgtgtgtgaaa 664
QY ggcttagtgtctcgatatacccaagaacatttgtgtcagaagctcttcttaaatgct 829
Db 665 taagtccagtaaccgaatgtcccaaggaacacccgtggcacaagatatgtgtcaagtgc 724
QY tatcactcaagtgtgaaatgtgaaattgaaacccatttctgcaagtttggtttatagt 889
Db 725 tgaaccttgaagttcgagatgaaattgagccctgttgcagacattgtccctcagatg 784
QY tcaaggaagaagaagaatttcagaaaaacttattttagacctaatctttagagacagtga 949
Db 785 ttaaggaagaagaagaatctccagaataatttcaactgtgacctgtgacctgtgacagatca 844
QY aaggtgtgtacgtccacatgtacacacgtgtgcacattactacccctgtgcaagatcagaa 1009
Db 845 aaggttctctgaggtcacaacgcttccagtgtgcgcgcacaaagtcagagcgagatcttcag 904
QY tttttctatacacttatccttcccaagatgttttctgttaataagctagaanaagctc 1069
Db 905 tcttctcagtcacacacccgtcctcagacatctactctgtagtaagattgaaaaagctc 964
QY 1070 tacagcaaggagacatttgagaggtgtgcaagaacatatatatttccaagaagacagt 1129
Db 965 tgcagcagggagatatgtggaacgtgtgcagagccctacacggtctatccaagaagtgatg 1024
QY 1130 ccaaccaagaataaagaanaaactgtgagaacactgaaagatcgaagcagatcaagtttgccaaa 1189
Db 1025 gtggaagaagtaagaagaagaattgaaataaactcaaacctcgaacgtcgaactccttcgcgcgc 1084
QY 1190 gactgtggaataatcgcagcttctgtctgtgactgtgcaatccatcttaatgaattgttta 1249
Db 1085 gtcttggtggaataatccggtgtcccttgccttgcgtgtgacccatagcttatacaagctcttc 1142
QY 1250 gcaagtcgtggtgagtttgaaagaagattctacagaagtgaatcaogtatcttgagaaagaa 1309
Db 1143 -caatgtcccccctctgtgagggaggttaactgtatgtgga-----ctctgtgtgtgga 1195
QY 1310 aaggtcttctgtcagagagggaggaattctatgtatgtgtgcagacgataccttgaaaaga 1369
Db 1196 gaagcccaagtggtgaaagggagacattgtgcaccaatctagaaggtcttc---tgaagaag 1252
QY 1370 caacaagtgtgagatgtatgtctgttaactgtgacgacttctgcagcagctactcttcacagtga 1429
Db 1253 ccccttctctgtgaggaataatgggtgtgtgatacctaactcaaaacccctcactgtgagctta 1312
QY 1430 caaatcttcttgaaggaagaagagcgcgttaagtgtgatgaatcttcaaaaattctctg 1489
Db 1313 gcaagcttcttcaagagaagaagatgcgccttaagcatgaaagacttatacaagtttttag 1372
QY 1490 ctgatatagaaagacacacttctgtcttactacggtcagactaagacctaatacactcagctca 1549
Db 1373 ctgactacaagaatcatcatccttactacagagatgcaggtcaagttccagcttgctcaa 1432
QY 1550 agatgacatacttccgcacactgtgaaatcccatattatgcctaacctccgagctgtctc 1609
Db 1433 gactgtgaaattcttaacactccagagatcatcaatgtgtctgtactcctcctaagatgtcgc 1492
QY 1610 aagtgaagcttaacacttaacagtagagttagacactacagagaanaatcttaagtttccg 1669
Db 1493 ccgtgaaaccccttctcctgaaaaccgacacgcccgcacaaagagattttagaatctccaa 1552

QY 1670 caaggatgtttagtcttccaacaactactacagaatactctctacataataccctcaga 1729
Db 1553 caagagagaatatagttccctccaacactgtgtacagaanaactctctatagtctaacccaaga 1612
QY 1730 gtcttaatttgcacatcgttcaaggttctgtcagaaataataacagtgtgaagctcgttta 1789
Db 1613 ggtctgaactgtttaaacaacttagcatcagcccggaacttacaataaagctcggttta 1672
QY 1790 tgtatgagagagatccaagaacatgcagctacgttaactctttagtaatcattagctgtcag 1849
Db 1673 tgtgtgagagagatgtcgaatgtcgaatgcagctcgttgcactctttagaaatccagcggtctg 1732
QY 1850 aatttcaagaagagcctatcacagcgttagtatatcatacagaagctcctgtatttctag 1909
Db 1733 aatttgcagagaagtgtaacacagcgttatacatataataaagtctcctgtactatg 1792
QY 1910 aagaatacgaagttaagctctcctgtcactttaactgtacgacatacaactgtcttttact 1699
Db 1793 aagaagtgaanaatlaagctcccgctaagctcacagtaataacacacctcgttcaact 1852
QY 1970 ttatcatgttagtgtlcaacaaaacaaataactcctcttgaacacacagtttgatata 2029
Db 1853 tctacatatcagctgttcaagcagaagaagagcctcgttgaaactcctcggtatatt 1912
QY 2030 catgtataccaatgtcttcagaaatgtgacggttgaagaatggtccagtttctgtgcagctc 2089
Db 1913 catgtgcacaaattctcttaaatgaacgttcttcaaaccttgatctactgtctccagttg 1972
QY 2090 catgtgaaaaaacacacagagctatctctgtactgtctcctgag-----gttccctc 2140
Db 1973 ccttgtgaaaatgtgccacccaactactcactgtcactctgtgagaagatccacattacaga 2032
QY 2141 tacctgtcagtaatgtgtatgaatcacaagaagtgltttaaagtgtgaagttgtgtcgtg 2200
Db 2033 atcctcccaatlaagtgtggtcgtgaagaacataaggtatattatattgaagtgcgaagctg 2092
QY 2201 ttctgtctatcatcatcacaagaatccttatcttgcacaatatttctgtcgtgcatagtctc 2260
Db 2093 ttctctgtgaacacacccaagacacacacctgtgagaagttcttccacccctcgtccactccc 2152
QY 2261 tggatgaacacctg-----tcccagctccgaattgtgggaatcgtgaatcaggaanaaact 2317
Db 2153 tggagagcaggtgtgaccttcccatccgcgtgtgtgtatagaanaatcagcggaatgtgcgc 2212
QY 2318 tagaanaatgaattgaaagagagatatttcaagcactgtatcatccagcgtggaacagatg 2377
Db 2213 tggagcattgtgcgtgaagctcagcatcatcgtcgtgaacccctccgcgtgtgagccgctcg 2272
QY 2378 tccgatttctcatctctgtctagataaactgtataactttagttatattagacctcctgtca 2437
Db 2273 tgccttctcgtcaacccgtgtgtgtgacagactcttccagctgtcgtgcagcccatgtgca 2332
QY 2438 ttgtctgtccaaatagtttaactagttcgaagcatctttagaagcagatgcataattataa 2497
Db 2333 tgcgtgtgcagacaagccaacttctccagatttgccttgagttccgtgtgtgtgcatagcaca 2392
QY 2498 atcgacttccaaaactgtgaaaggaatcatgtaccagcagctgtgcagaagaacgtctcttg 2557
Db 2393 acagtcgcacaagaagaagacctgtgacgaagaacagcatgtggaagaactgcgtctgtg 2452
QY 2558 catcataatattatagtlttccgcctacccaataacttaacctaat---tcatcatac 2614
Db 2453 ctctcactagcgaactagcttctcgcctgcagaggtgtcagaaggtgtgtgccaaagtcag 2512
QY 2615 caggtcctgtgggtgttgggtgagatcaggtcatattgtccaaatagtgctagatcgtgtgta 2674
Db 2513 ggcgtcccaatgcctctcagacctcctogagctacacacaagtatgtccgacatcagctg 2572
QY 2675 gacctcgaagccttaatttaactcgttctcgaagccttgaataagaaatccagatatat 2734
Db 2573 ctgctgtgagattcaaaactgtctgcagcccggtgtgattgagcagcagtaaacccagactctg 2632

QY	2755	ctgagactcccaagctcccaagatgagatgaaatttcgaatcatcctggagagtaagggtttag	27594
Db	2633	cgggagacacactcccgacagcagcagggagaaattgaaagacatagtcttccaaatgcgcg	26322
QY	2795	atcgctccaattcctctggtttaacactggtgtgtccaaagctgccccatggtgatccaac	2854
Db	2693	atcgcaactgagc-----	2706
QY	2855	ccagatccaagtgccagaaatcaacacaggtatggtatcgaaattgtatcgtgtcttcgc	2914
Db	2707	-----	2731
QY	2915	acacagagacgtcaagattcttcaacaaatcaacgagacgtctaccacataaagctt	2974
Db	2732	gtatgtgatgtgtcccaagtt-----	2785
QY	2975	ttccaggaagacgtcgcttgcagtggtgtttgcaggtgtgcagcgcttcggaaatcagctt	3034
Db	2766	tccatgagagagctgtgccttccagatggtgtgtgcagccgcg-----	2825
QY	3035	tgcacaagcctgtgtcttttttgaatlaagtgaaagacatggtgcacattatatt	3094
Db	2826	-----	2857
QY	3095	ttaatgataaactctgaggtccccaaggaagaagtcgtttccagaacggttcaatgtatga	3154
Db	2858	acatgagcaacaacgggacacagttcccgagagagactggttttctgaacggttcaatgtatga	2917
QY	3155	ttgcagcctctgttcagacagatgtgcgtatgatagtttccaagatt-----	3205
Db	2918	taactactattgttaatgvtgtacactcggaaattgtgcagccctttagtaaacccaaga	2977
QY	3206	aggaacacagaatgtgtgagagagacataacaaagccttgcattcttccatgatcgt	3265
Db	2978	aggaagaatgagacagcgagaaagaatgaacatacgaactgcgtcttctctctgtatgaacttc	3037
QY	3266	tgctgtgtatgagacagagaattgtgttttaagccctataaagtcctgtcataacagatgt	3325
Db	3038	tctcccatatgagatcgggggtctgtgttttaaccctacacagatattatgtgcagcagctgt	3097
QY	3326	cttcaagacttactcatctatccgaatcccgacggtgtcgtgtgccttgggtgtgatcttc	3385
Db	3098	cagccaagctccagtaaccttccaaagctca-----	3148
QY	3386	tacgaatcatctgcagatgcagagacagatgtatcatataaacttaaccctgcagcttactta	3445
Db	3149	tgaagatccctctgttagcagatgagcattactccaatctgaact-----	3202
QY	3446	ctccaactgcatactcatcaactcctctgtttcttctgcgaacatctcagaatctcgaattt	3505
Db	3203	atgtctgatactgtcccaacatactctctgtccttgcatacttcccaagaactcaagctcct	3262
QY	3506	ctacgaatgtacaagaccaaaagatgtgcgaatattgttgaattacgcgtgccttcgcgc	3565
Db	3263	gtctcagacttcccgagaccagaagaatctgcgaagaatgttcgaatcctgcagatcagcc	3222
QY	3566	aacagcatatttgcgcagagactgtgtttaaaccagagctgcgtgtcaattttaagacctatg	3625
Db	3323	agcagcacttccccaacccggctcctctcttcaacagaactgcgtgcgtccctgtatgcgaag	3382
QY	3626	ctgaagagactggtttgattgtgacataaagaatcaataatgtaacaaattactctcca	3685
Db	3383	ggagagagaaatcagacaagatatacaaaagagcgttcaatgacatltcaacagccttgtagt	3442
QY	3686	gtccagcgtccagcccgcggtactctgcaccctcagataaagctccggtgtgcattgtgt	3745
Db	3443	ctccagcctccggaccacagcctgtgtcaaacccagaggtgaagttcaaaatgcgcgcctt	3502
QY	3746	atctacactctgattgttattatcatatgaagaaactgtaccccaagctgtatatttttagaa	3805
Db	3503	acctacacttagttagtcatcatatttggatgacttgcacagcctctgtgactttacagtgtg	3562
QY	3806	ctcaacaacaagaggaagaccaaatttgatgtatgacacatgtatattatgaagtgaagcg	3865

Db	3563	caagatctcgcaagataccgacacagctgctcg-----gatgaagacagaagagcg	3613
QY	3866	gaagttgtaagccagagacgtttgcacatggacatgcagagagacatcgtgctccaactaa	3925
Db	3614	ccggtgcacataaccagaaatgtggtctctgctgcacatagccagggataattccaattgnaaa	3673
QY	3926	caagcgctgcagattccctccctcaacgtcccaacagatgvcagggacaacacactactcttcag	3985
Db	3674	caagtg-----gaatagtgctgtctctcttgccatataagcagatcaacacatgctgcagcg	3727
QY	3986	cagaatcaagtcogaagccttttgatctgtctctactcttgggtcttccaanaatgycagatgaa	4045
Db	3728	cggacactactcgcaaccctcagatctgtctctctctgtgatacagtaaaaaatgctgatacga	3787
QY	4046	cagttctaagaagtggtttacagatctctcctgcttgcagctaaacggctatagatc	4105
Db	3788	ggcctaataaggatgagattgtctgacccctcccaacgaagctccaacagatattttagatc	3847
QY	4106	tgtttatctctgtgtctgtctgtcttctgagataaaggaggaanaaagtgtttgaagcaatga	4165
Db	3848	tactttcaactctgtgtgtatgttttttgatataaagggaanaaagattctgcacaagtcca	3907
QY	4166	atagtctgacctttaaagaaatcaanaagacatgagagcaaaagcttgaagagactattctg	4225
Db	3908	gtaccacaagctctgcagaaagctaaaggtatgataagcccgctgtgaagaggtttgtctgc	3967
QY	4226	ggaagcattagtgccagcgcaagaatagtatccgctgaagccggagagacagctccgagaaagcc	4285
Db	3968	gttggaagaggccagagaggagatgatgtgcgcgcggctcccaaggaaacggcagatttc	4027
QY	4286	catctggaagtgctcttggaagtcagaanaatttgagtgvgagaggaagaatatgtaact	4345
Db	4028	ca-----ggcctaaatgnaaatttgatgagtggaagaaagagcgcagacacatt	4072
QY	4346	ggcgttcaaaacacagagaagccttgacaatatcaagaagcagagatltgaacacgaagactga	4405
Db	4073	ggcggcaagctaatgagaagctagatataaacaagaagccgaagttagatcaagaagccttga	4132
QY	4406	ttagtgaaaccttgcttatacagaagaacactaatattttagatacatattagatattgtg	4465
Db	4133	tcaagtggcaactctggtctatacagaagacattatcatatccctgtgatacgaagaaacattaa	4192
QY	4466	ttcagaacgcttctgttaaacggaatcccaagaagagacattctgtgtgagtgctcaaaagtgc	4525
Db	4193	tccagcgagcctcggtctcgtgcactgttaagaagaagcctgctgvgaggtgttctcgaggtgc	4252
QY	4526	tactacaacagcatgycgtgttaacaaagatgcaatttactatacaacatggtttgtgtacac	4585
Db	4253	tgtgtaattctctgaactgtgatacagatgataccactactcgtacacgtcttggacaacac	4312
QY	4586	agaagcgcttggtttccaagtttctctgaactctatttgaagaagaagacagagcgtgtg	4645
Db	4313	tccgtgcttccatccgcacaaagttttgtagacttaacttccgaagaagaggtggaagtgct	4372
QY	4646	ctgatttaatgcctcagagcttctccgaacactgtagcagtagcacatgctgtccaatacagctac	4705
Db	4373	tcgacttaatgataccaagaagccctgcacacactgtagcagtagcacatgatactcaccggagcc	4432
QY	4706	accocagtgctcctcttaactactaataatgagcaaaacttggagatltgggaataacttgg	4765
Db	4433	aagcctgtgccacccttataccctccatagaggttcaagttttggagccacagtaatttgg	4492
QY	4766	ccaggtgttaaaatgtaggttacaatagtcataactctcctctgtgtggcgacatctcgaatt	4825
Db	4493	caaggttaaaatgtaggttacaatagtcataactctcctctgtgtgtggagagcaccagact	4552
QY	4826	ttaagaagaatcttcaagacgtttctcaagaagacataatgacatagctgaagaagatc	4885
Db	4553	ttaagaagaagacactgtgagagatccttgagagacaatttggcctattcagaagaagagaca	4612
QY	4886	tggaaattgagggaaacacatttctctgtatcaggttcagagatctggttttcaattccaata	4945

Db 4613 caagccatgcatgactcctttccaccacagtgaggaaacttctctgtaatctgtaata 4672
 QY 4946 tgattcttcggtaactgtgtaaaatgaagaaacacaggaagatcttaaatgtgtatg 5005
 Db 4673 gcaacttatatgaacacagtgtaaaatgaaggaaatctcaagaagatctgatatg 4732
 QY 5006 atctaattgacagaattgcacaagggtaccagaactctccaga--gcgattgacctg 5062
 Db 4733 attcattgtaacagaattgcacaagggtaccagaactctctgattctgctgacctg 4792
 QY 5063 tgcagaacatgagcaggaacactctccagaacgaagcaatcatgcttaagtctcacaggtc 5122
 Db 4793 tccagaacatgagcaggaacacacacagaagaagtgtctacacaggaagctgcacatg 4852
 QY 5123 tagtcctcagcagcagcctgtgtctgtaaatattgagcaatgctgaggaagcaggaata 5182
 Db 4853 tgggtgcacagcgcgtctgtagtgcgtgatatctgagcagtctgtaggacacagctacc 4912
 QY 5183 ttctctgaggatgtgttaactcttcagaataattcaattcaattgttttaagaagatctgc 5242
 Db 4913 tgcctgtgagcagtgtagctctccagaataattctctcagaatgtctgtaggaaggtctg 4972
 QY 5243 tctcagaatgtagtctctccagaatgaagatgtaatctgctctgtaaaataacttactg 5302
 Db 4973 tctctgaggaacacctgtaccctgcagcagagatgggtgtgagcaggaactctaccg 5032
 QY 5303 agtcaagaactgtgaggatctctgaaacagcagctgtctctctctatgctgcatgt 5362
 Db 5033 agagtgcctgtgtagcctctgcagcagcgcgagagctctctccagcaggaagctcat 5092
 QY 5363 atgaagcagtaatgaagttacaagttacttaattcttaattgaagctaacggaatg 5422
 Db 5093 atgagacagtaatgagtgcttacaagctgtgtaactcccaactcctagaagcgatcgag 5152
 QY 5423 caaagaactatccacacattcatgtaactgaactcaagaacatccagcaaatgtctc 5482
 Db 5153 tccggaacatgacactaccacacagcagcgtcgaagacctctccagcgttaca 5212
 QY 5483 agagtactgctgagcagcagatgttggcacactattctgtgtgttttaagaaaca 5542
 Db 5213 agatcatc-----aagagaatgttttgaaactactctcgagtggtttcttctgatacca 5266
 QY 5543 agtccggagatttgatgaacaagaattgttttaacaagaagcctgacaataaccaactg 5602
 Db 5267 aattcggagatctgagtaacagagagttgttccacaagaagcctcgaattaccaagctc 5336
 QY 5603 cagaatatctcacaagattgagagatttaacggaagaagatttgaagaagatgtgtg 5662
 Db 5327 ctgagatctcacatagactagcaggaatttatgttcaatggtttgtgagagattgtg 5386
 QY 5663 aagtatacaagaacttaactcctgttagacaagtgttaaatatagatccatacaagcata 5722
 Db 5387 aagtattaaagaactcactcctgttgacaacacaaagttgattcctaacaaggctaca 5446
 QY 5723 ttcaattactatgtagagccataacttgaacacatagatgaagaacagatccact 5782
 Db 5447 taagatcactcttgtgtgagcctacttgtatgatatgataagtaagaagcagatccat 5506
 QY 5783 atttcgacaaaatatacaactcctcgttagatcaatgtatcaatccacttacttaatg 5842
 Db 5507 acttggaaagaattcaactcctcgagagttcatgtacacaccccggttaccacccgag 5566
 QY 5843 gccgtgccatgggaaacttcatgaacaactcaagaaggaacccattgactatgctc 5902
 Db 5567 ggcgcgcctcgaggagctgtatgagcagtaagaagaacacagctccgacacataagc 5626
 QY 5903 atgcttctcttataataaacaagggtcaatgtcaactataagaagatcatcctaa 5962
 Db 5627 acgcttctccatcatacaagacagcagctcagcgctcacaggaagagaggttgttga 5686
 QY 5963 caaccaatgaagtgtcatgttagagacatgcaagaagaacacaggaaggtgcatgcaa 6022
 Db 5687 caacgattgaagttgcattgaagacatgaagaagaacacctgcagtttagcagttgcca 5746

QY 6023 caactcaagatcccgacagcccaaaaatgtctcagatggtagtactccagagatctgagca 6082
 Db 5747 ttaaccaggaagccgcctcgatgacaaagatgtcttaagatgtgtctgcaagcctctggga 5806
 QY 6083 ccacagtgtaacaggggaccttggaaagtgtgccaggttttctctcgtgaataactagtg 6142
 Db 5807 ctaccgttaactcaggaagccctggaagtaagcccaaggttcttggctgaataatctcgtg 5866
 QY 6143 acccaagccttcaagacatcataataaactgcatctgactcgtctttaaagatttactaaa 6202
 Db 5867 atccaaaactctacgcatacacaacaaagtgtgtagtactgctttaaagatcatcagta 5926
 QY 6203 ggtgtgaagatgtccttaagaanaataagaacttaattggtccggttcaaaaggatctc 6262
 Db 5927 gattgtgtgaagctgtatgagaanaacaaagctctcatcagcagacagaggaataac 5986
 QY 6263 aaagggaattggga--aactatctgccttaagaagccctacaagccttagatc--ac 6318
 Db 5987 agcaggaactcaaaaagacatacaagaactcaaaagagaaccccaagccaatgtacgagc 6046
 QY 6319 aagaagctccagttatccaaagccaatatt 6349
 Db 6047 ggaaaattccagaactgtacaaagccaatatt 6077

 RESULT 8
 AAS07382
 ID AAS07382 standard; cDNA; 4027 BP.
 AC AAS07382;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE Human DNA associated with CLASP-5 #2.
 XX
 KW Human; CLASP-5; cadherin-like asymmetry protein; immune gateway;
 KW Immunogen; antibody; autoimmune disease; rheumatoid arthritis;
 KW Multiple sclerosis; leukaemia; insulin dependent diabetes mellitus;
 KW acquired immunodeficiency syndrome; AIDS; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH CDS 1..2970
 FT /*tag= a
 FT /product= "protein as displayed in AU004026"
 FT /partial
 FT /note= "No start codon"
 XX
 PN WO200142296-A2.
 XX
 PD 14-JUN-2001.
 XX
 PF 13-DEC-2000; 2000WO-US34163.
 XX
 PR 13-DEC-1999; 99US-0170453.
 PR 14-JAN-2000; 2000US-0176195.
 PR 14-FEB-2000; 2000US-0182296.
 PR 11-APR-2000; 2000US-0196267.
 PR 11-APR-2000; 2000US-0196460.
 PR 11-APR-2000; 2000US-0196527.
 PR 11-APR-2000; 2000US-0196528.
 PR 11-APR-2000; 2000US-0547276.
 PR 13-OCT-2000; 2000US-0240503.
 PR 13-OCT-2000; 2000US-0240508.
 PR 13-OCT-2000; 2000US-0240543.
 PR 13-OCT-2000; 2000US-0240539.
 XX
 PA (ARBO-) ARBOR VITA CORP.
 XX
 LU Lu P, Garman JD, Candia AF;
 XX

DR MPI: 2001-367865/38.
DR P-PSDB: AAU04026.
XX CLASP-5 polynucleotides, proteins and antibodies are used to prevent or
PT treat a CLASP-5 mediated disease, such as an autoimmune disease e.g.
PT Rheumatoid arthritis -
XX
PS Disclosure: Fig 9C; 188pp; English.
XX
CC The sequence encodes a protein presented as AAU04026, included in the
CC specification which relates to a cadherin-like asymmetry protein,
CC CLASP-5, which is a transmembrane protein of the immune system involved
CC in the formation of the immune gateway. CLASP-5, polynucleotides encoding
CC it and an anti-CLASP-5 antibody are used to prevent or treat a CLASP-5
CC mediated disease, such as an autoimmune disease caused or exacerbated by
CC increased activity of TH1 cells. These diseases may include Addison's
CC disease, rheumatoid arthritis, dermatitis, Multiple Sclerosis, Systemic
CC Lupus Erythematosus and Autoimmune Thyroiditis, inflammatory conditions
CC (e.g. Ischaemia-reperfusion) and responses, Leukaemia, acquired
CC Immunodeficiency syndrome, AIDS, syphilis, typhoid, Burkitt's lymphoma,
CC Grave's disease and insulin dependent diabetes mellitus. CLASP-5 is used
CC for inhibiting an immune response in a cell (T cell or B cell) by
CC interfering with the expression of a CLASP-5 gene in the cell, the
CC ability of a CLASP-5 protein to bind to another cell or the ability of a
CC CLASP-5 protein to bind to another protein. The antibody binding CLASP-5
CC is used to inhibit an immune response is a subject. The polynucleotide is
CC used to detect CLASP-5 expression in cells and for diagnosis of
CC diseases and disorders associated with aberrant expression of CLASP-5.
CC Note: The present sequence in the specification but is
CC not mentioned anywhere in the specification.
XX
SQ Sequence 4027 BP; 1162 A; 898 C; 957 G; 1010 T; 0 other;

Query Match 15.8%; Score 1079; DB 22; Length 4027;
Best Local Similarity 62.9%; Pred. No. 2,7e-261;
Matches 1836; Conservative 0; Mismatches 1035; Indels 46; Gaps 9;

QY 3443 ttactcaactgcatctccatccactctgttcttctctgcaacatccaggttctgagt 3502
DB 8 tgaatcgtatagatcgtccaaacatctctgtctcttccatattctccagaaatcaagt 67
QY 3503 ttctcagaatgctacagaacagaagtcaaatatggttgaatt--atccgtgacctt 3560
DB 68 cctgtctcagcttccagagcagaagaatgcccagatgcttcgactcgacttccagagtgac 127
QY 3561 ccgcacaacagca-ttatttggcagagactgtgttacaagacgtgctgtcatattttagacc 3619
DB 128 ccgcagcagcagcacttccctcaccgcgggtctctctctcacaagacatgctgtgccttgatg 187
QY 3620 ctgattgtcgaaggaactgttttgattgataagaagatcacaatattgatacaatttac 3679
DB 188 ccgaagggggaaggaatcgcgaagtacaagaagactgtccaagcaattccaagacctgc 247
QY 3680 tctccagttcagactcagaccgcggtactgtaccctccagatataaaggtccgagtgcca 3739
DB 248 taagtcttccagacttgaaacccaagcgtgtgttcaaacagaggtgaaggttaaatccgcg 307
QY 3740 tgttgtatctactctgtattgtattatcatatgaaactgttaccctcagctgtatgattta 3799
DB 308 cccttactactactttagttgtgcatcattttgtgagtcttggccacaagctcgttgatac 367
QY 3800 ccgaactatcacatataaagaggaagaacaatttgtatagccactgactgatatataaagt 3859
DB 368 ccagttgcagatatctgcagataccgcaccagtgc-----tcggatgaagaataaag 418
QY 3860 agagcgaagatactgaacagcagacctgtgccatgtgcaatgcagcaggaatcvgttccctc 3919
DB 419 aaggaagccggtgcatctaccagaatgtgctgcgcacatagcaggaataatattcaatt 478
QY 3920 aactacaagccctgagatttccctcccaagtcacagagtgagcaggaanaaacacttacc 3979
DB 479 tgaataacagtg-----gaatagtgctgtcttcccttgccctataagcagtaacaatgc 532

QY 3980 ttccagcaaatcaagtcgaagccttttgatctctactcttgggttcttcaaaaatgcag 4039
DB 533 tgaacgcgacactactccgaaccctcacatgactcttcttccttgatataatgaatgtcg 592
QY 4040 atgaacaacttctacagaagtggtttacagatctctcagctcttccacataaacccgctat 4099
DB 593 atcagagcccatlaagaagtgaattgtctgactgcacatcaacgagctccaacagatctt 652
QY 4100 tagatctgtatctatctgtgtgttcttgcctttgagatataaaggaaaaagtgtttaac 4159
DB 653 tagatctacttctacatctgtgttattgtttgagataaaggaaacaagaattctgaca 712
QY 4160 gaatgaatgcttgacctttaaagaataaagaacatgaagcagaagcttgaagaacta 4219
DB 713 aatctcagttcccaagctccgcagagttcaaggatgtaagcccggtctgtaagaagctc 772
QY 4220 ttcttggagcataggtgtgcgaagaatgttaccgagcgaagccgaagcagctcgaga 4279
DB 773 tgcgtgcgtgggaagggcgccagagagatgattgctgcgcggcgtcccaaggagacc 832
QY 4280 gaagcccatctggaagtgctccttggaaatcagaagaatttgaagtgagggaagataga 4339
DB 833 gatttcca-----ggcctaattgaatattgagatvggaagaagagcaga 877
QY 4340 ctactgtggttcaaaaacacagaagaacttgacaatatcaagaacagagattgaacagag 4399
DB 878 cacattgtgcgacagctaatatgaaagctgataataaacaaggccgagttagatcaagag 937
QY 4400 cactgatttgaanaacctgtgtacagaagaacctaactcaattttagatattagataga 4459
DB 938 ccttgatcagtggaatctgtgtacagaagcacatttcaatctctgtgataatgacagaa 997
QY 4460 ttgttgtcgaagcgttttctgtaacgaatccaagaagacattctgtgtgagtgtcaa 4519
DB 998 acattatccagggcagagctgcgtgcgtgactgttaagaacagcctgtctggaggtgtcga 1057
QY 4520 aagtgctactacaacagatgtgcgttgaacaaagtgcagttatatacaaacgtgtttg 4579
DB 1058 ggggtgcgtggaattctctgaactgtgatacagagttaccacttccactgtactcgttgg 1117
QY 4580 ctacacagagagccttggtttcaaatgttcttgaactcttatttgaagaagagacagagc 4639
DB 1118 caaacactcgtgtctctatctacgtcaagtttggagacttacccttcgaaaggagggtggaac 1177
QY 4640 agtgtgtgatttatgtcctcagagcttcccgacatgttagcagttagcatcgtgtacaatac 4699
DB 1178 agtgttccgacctatgtccaccaagttcctgtgacacatgcagcagcatgagtgtaaccc 1237
QY 4700 ggtcacacccagtgctcccttaccctactaatgagggcaaaactttgagattggaata 4759
DB 1238 ggaagcgaagcctgtgcacaccttaccctctcctatgaggttcaagtttggagccaacagta 1297
QY 4760 actttgcagaaggttaaatatgaggttaccatgttccactatctctcctgtgtgggacatctc 4819
DB 1298 attttgcaagagttaaagttaacatgtaacatgttccctgtgacatttgggtgggaagcacc 1357
QY 4820 agaatttaagaagaattcttaagagcttctcctaagaagctatattgacatattgtgtaag 4879
DB 1358 cagactttaatgaagagagcacttgagaagatcccttgagaacaatttggccattacagaag 1417
QY 4880 aagatctggaattgaaggaaacaacatttctcgtatcaggtccagggatctgtgttcaacc 4939
DB 1418 aggaacacagcatgcaagatgactcttctccaccacgggtggaggaactctctctgttaac 1477
QY 4940 tccatgatcttcttctgatactgtgaanaatgaagaacacagggaggaatctgtaagt 4999
DB 1478 tgaataagcatcttatatgacacagtgtaaaatgaggaatttcaaggaaagcctcgagatgc 1537
QY 5000 tgattgatctaaatgtacagaatttgcgaaggtttacagacctctcaga---gcgattga 5056
DB 1538 ttatgatctcatgtacgaagattgcgaaggtttacacagcatctcctgtactcgtcggtcga 1597

QY	5057	ccgtgttcgcaagaacatgugcagcgcaagcactctcagaaacgaagcaatcaatgctcgaagctgcac	5116
Db	1598	ccctgtccctccagaaacatgycagagaaacacacacaaagaagaagtgtctacacggaagctgcaca	1657
QY	5117	agttctcaagccactcaagccagccactgtttgttcgtgaatatttggacatgtctgtggaacgcgga	5176
Db	1658	tggtcccggtgcaacgacgcgcgtgttgagctgcagatgatactgagcactgtctcggaaagaccaca	1717
QY	5177	aatactctctgtggtgattgtgttaacattccagaaatatattctcaattctttaaagaagat	5236
Db	1718	gctaacctgcgcggtggtgagagtgcatgcttccagaataattcttccaaatgtgctctggaagagt	1777
QY	5237	ctgcgcgtctcaagatgatalgtgtatctctccagatgaaagaagatcatctgctctcgaaataact	5296
Db	1778	ctgtgtctctgtgaggaacacccctgtccacctgcagagatgggtgtgctgcagagccagtaact	1837
QY	5297	ttaactgtctcaagagactgttggtggtattctgtgaacaaagcagctgtcttccttctcatgtctg	5356
Db	1838	tcacccagagagtgtgcctgtgtatgagctctctctcggagagccgcggagctctcttcagcaagga	1897
QY	5357	gcacgtatgaagcaggtttaaagaaagtttacaagaagtaacttatctcatatgaagtaact	5416
Db	1898	gcttatatgagacagttaatgaggtgtctacaagctgtgtcatcccatctcctagaagcgatc	1957
QY	5417	gggatgcacaagaacactatccacaacattcatgtgtaaacattcaagaagcattcagcaaaaatg	5476
Db	1958	gagaaattccggaagctgtacacactctacacacagacaagctgcagagagccttctgcagcagatc	2017
QY	5477	tctacagagagactgctgttggaagcagatgttttggccactatttcgttgtgttttatg	5536
Db	2018	ttacaagaagacat-----aagaagaatgtttggaacactcactccagatgtgttctcttg	2071
QY	5537	gaacaagcttcgggattttgatatgaacaagaattttttacaagagagcctgcataacaca	5596
Db	2072	gattccaattggggattttgtgatgaaacagagatttgtctacaagaagcctgcataacaca	2131
QY	5597	aacttgcagagatatctcacagattggaaggaatttacaagaaagaatttggaaagagatg	5656
Db	2132	agcttcctgtgagatctccacatagactagagcattttatgtgtaactgtttgttgagagat	2191
QY	5657	tggtttgaagtaataaagaactctcaatccctgttagacaagtgttaatatgatacttaacaag	5716
Db	2192	ttgtggaaggttaataaagactccactctctgtggaacaacacagatttggatccttaacaag	2251
QY	5717	catatatcagattacactatgtggaagcgaacttgcacatacatatgaattggaaggaacaa	5776
Db	2252	cctacatacgaatacactttgtgtggaagccctactcttgaatgagtatgagaatgaaagaacagg	2311
QY	5777	tcacactatltogaaaaaattataacattcttgcctatgatactgatactgcatacccttaact	5836
Db	2312	tcacatacttttgagaagaatttcaacctccgcgggtgtatcgttaacacacacccggttaccac	2371
QY	5837	tagatgacggtgcccatggtgggaacttcatgagaacaattcaaaaaagaagaccattctgacta	5896
Db	2372	tggaagtggtgcgcctccggtggagagctgcatagtagaagatagacaagaagctccgtgacca	2431
QY	5897	gcctccacagcctttccctatatattaacaaaggttcaatgtccactatataaagaagaagatca	5956
Db	2432	ctatgacagccttccctcaatacaagaacagagctcagctgtcatccagaagagggaggtttg	2491
QY	5957	tccttaacacccaattgaagtgtctattgtgagagactgacaagaaaaaagacagagagttggcat	6016
Db	2492	ttttgacacgagattggaagtgttcattgtgaagacatgagaagaagaccctgcagtttagaag	2551
QY	6017	tttgcacaacatcagatctccgcagacaccccaaaatgtcttatagatgtgaatccagggatctg	6076
Db	2552	tttgcattaaacacagagagccgcctgtatgacaagaagtgtcctacaagtgtgtcgtcaagctctg	2611
QY	6077	taggacacaacagtgaattcaaggggccttttggaaatttccccaggttttctcgtctgaataac	6136
Db	2612	ttggagagctaacgttaataccaagaggaacactgtgaaatgaagcccaagtgcttttcttgctgaatacc	2671
QY	6137	ctagtgcacccaagctctttagacatcatatataaacatgtgcactctgtctttaaagaattttaa	6196

Db	2672	ctgctgcatccaaaac	ctcatalcgacatccacaacaa	ggttgagtgtaagcttaaggaattca	2731
Oy	6197	ctaaaggtgtgaagatgccttaaga	aaaaataagagcttaattggccggttca	aaag	6256
Db	2732	tcatgagatgtgtgtgaagcctgt	gagaaaaaacaagcgtctcatalc	acagcagacccaaggg	2791
Oy	6257	agatacaaaaggaattg	gggga--aactatcttcgcttaaga	agagccctacagccctaga	6314
Db	2792	aatacgcaggaactcaaaa	aagaactatacaagctaaagaga	gaacccctcagcgccaatga	2851
Oy	6315	tc--acagaaagtcctc	cagttatccaagccagtaatt	6349	
Db	2852	tcgagcggaaaatcc	cagaaactgtlacaagccaatatt	2888	
RESULT	9				
AAH43850	ID	AAH43850 standard; cDNA; 2148 BP.			
XX	AC	AAH43850;			
XX	DT	04-SEP-2001 (first entry)			
DE	XX	Preliminary human CLASP-7 encoding cDNA sequence.			
KW	XX	Human; CLASP-7; cadherin-like asymmetry protein; immune response;			
KW	XX	neuroprotective; antidiabetic; immunosuppressive; antirheumatic;			
KW	XX	antiarrhythmic; hypotensive; anti-HIV; cytostatic; immunostimulant;			
KW	XX	antihaemic; antiinflammatory; ophthalmological; nephrotrophic;			
KW	XX	antitumorid; antiasthmatic; antiallergic; antibacterial; gene therapy;			
KW	XX	chromosome 19q13.2; autoimmune disease; multiple sclerosis; toxemia;			
KW	XX	juvenile diabetes; rheumatoid arthritis; pruritic urticarial papule;			
KW	XX	hypertension; Rh incompatibility; ss.			
OS	XX	Homo sapiens.			
FT	Key	Location/Qualifiers			
FT	CDS	1..1932			
FT	FT	/*tag= a			
FT	FT	/partial			
FT	FT	/product= "Preliminary CLASP-7 protein"			
XX	XX	/note= "no start codon given"			
PN	XX	WO200142295-A2.			
PD	XX	14-JUN-2001.			
PF	XX	13-DEC-2000; 2000MO-US34152.			
PR	XX	13-DEC-1999; 99US-0170453.			
PR	XX	14-JAN-2000; 2000US-0176195.			
PR	XX	14-FEB-2000; 2000US-0182296.			
PR	XX	11-APR-2000; 2000US-0196267.			
PR	XX	11-APR-2000; 2000US-0196460.			
PR	XX	11-APR-2000; 2000US-0196527.			
PR	XX	11-APR-2000; 2000US-0196528.			
PR	XX	11-APR-2000; 2000US-0547226.			
PR	XX	13-OCT-2000; 2000US-0240503.			
PR	XX	13-OCT-2000; 2000US-0240508.			
PR	XX	13-OCT-2000; 2000US-0240539.			
PR	XX	13-OCT-2000; 2000US-0240543.			
PA	XX	(ARBO-) ARBOR VITA CORP.			
PI	XX	Lu P, Garman JD, Candia AF;			
XX	XX	WPI; 2001-381641/40.			
DR	XX	P-PSDB; AAB99540.			
PT	XX	Novel cadherin-like asymmetry protein-7 and polynucleotides encoding			
PT	XX	the polypeptide, useful for treating autoimmune disease,			
PT	XX	hypersensitivity, preventing transplant rejection by modulating immune			


```
QY 6161 atcataaactcgcactcgtcttaagaattactaaaaggtcgaagtccttaa 6220
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1685 atcaacaactcgcctcgtcttaagaattccttcaagaagaatcgtgagatcgtcgc 1744
QY 6221 gaaaaaataagcttaattcgtgcccgtctcaaaaaggatcacaaggaaatg--gggaa 6278
    | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1745 ggaataataagcctcgtcttggccgcgacccaagaagatccaccgtgagcttgagcgca 1804
QY 6279 actcttcgccttaaaaggccctacagccctagatcacagaagtccttcagtattcc 6338
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1805 actactgcgcctcgtcggagagctctgcagccctcgtcttaaccagcgcctgcctcagctga 1864
QY 6339 aagccagatctgtctgtccctcgcacagagatctcctcagtcgagatcgttcgcaaa 6398
    || || || || || || || || || || || || || || || || || || || || ||
Db 1865 tggcacccacccccc---accgcgctcaggaacccttgaacagagacgaagttccgaaag 1920
QY 6399 atgagctcttaac 6412
    || || || || || || || || || || || || || || || || || || || || ||
Db 1921 gcagagctctgagc 1934

RESULT 10
AAS07381
ID AAS07381 standard; cDNA; 2014 BP.
XX
AC AAS07381;
XX
DT 26-SEP-2001 (first entry)
XX
DE Human DNA associated with CLASP-5 #1.
XX
KW Human; CLASP-5; cadherin-like asymmetry protein; immune gateway;
KW immunogen; antibody; autoimmune disease; rheumatoid arthritis;
KW multiple sclerosis; leukaemia; insulin dependent diabetes mellitus;
KW acquired immunodeficiency syndrome; AIDS; ss.
XX
OS Homo sapiens.
XX
FT Location/Qualifiers
FT CDS
FT 2..2014
FT /tag= a
FT /product= "Protein as displayed in AAU04025"
FT /partial
FT /note= "No start codon"
FT /transl_except= (pos:1523..1525,aa:Arg)
FT /transl_except= (pos:1883..1885,aa:Ile)
XX
FN WO200142296-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-US34163.
XX
PR 13-DEC-1999; 990US-0170453.
XX
PR 14-JAN-2000; 2000US-0176195.
XX
PR 14-FEB-2000; 2000US-0182296.
XX
PR 11-APR-2000; 2000US-0196267.
XX
PR 11-APR-2000; 2000US-0196460.
XX
PR 11-APR-2000; 2000US-0196527.
XX
PR 11-APR-2000; 2000US-0196528.
XX
PR 11-APR-2000; 2000US-0547276.
XX
PR 13-OCT-2000; 2000US-0240503.
XX
PR 13-OCT-2000; 2000US-0240508.
XX
PR 13-OCT-2000; 2000US-0240543.
XX
PR 13-OCT-2000; 2000US-0240539.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu P, Garman JD, Candia AF;
XX
DR WPI: 2001-367865/38.
DR P-PSDB: AAU04025.
```

```
XX CLASP-5 polynucleotides, proteins and antibodies are used to prevent or
PT treat a CLASP-5 mediated disease, such as an autoimmune disease e.g.
PT Rheumatoid arthritis -
XX
PS Disclosure: Fig 9A: 188pp; English.
XX
CC The sequence encodes a protein presented as AAU04025, included in the
CC specification which relates to a cadherin-like asymmetry protein,
CC CLASP-5, which is a transmembrane protein of the immune system involved
CC in the formation of the immune gateway. CLASP-5, polynucleotides encoding
CC it and an anti-CLASP-5 antibody are used to prevent or treat a CLASP-5
CC mediated disease, such as an autoimmune disease caused or exacerbated by
CC increased activity of TH1 cells. These diseases may include Addison's
CC disease, rheumatoid arthritis, dermatitis, Multiple Sclerosis, Systemic
CC Lupus Erythematosus and Autoimmune Rhyoiditis, inflammatory conditions
CC (e.g. ischaemia-reperfusion) and responses, leukaemia, acquired
CC immunodeficiency syndrome, AIDS, syphilis, typhoid, Burkitt's lymphoma,
CC Grave's disease and insulin dependent diabetes mellitus. CLASP-5 is used
CC for inhibiting an immune response in a cell (T cell or B cell) by
CC interfering with the expression of a CLASP-5 gene in the cell, the
CC ability of a CLASP-5 protein to bind to another cell or the ability of a
CC CLASP-5 protein to bind to another protein. The antibody binding CLASP-5
CC is used to inhibit an immune response is a subject. The polynucleotide is
CC used to detect CLASP-5 expression in cells and for diagnosis of
CC diseases and disorders associated with aberrant expression of CLASP-5.
CC Note: The present sequence is included in the specification but is
CC not mentioned anywhere in the specification.
XX
SQ Sequence 2014 BP; 565 A; 469 C; 514 G; 466 T; 0 other;

Query Match 12.3%; Score 841.6; DB 22; Length 2014;
Best Local Similarity 66.2%; Pred. No. 1.3e-201;
Matches 1278; Conservative 0; Mismatches 639; Indels 13; Gaps 4;

QY 4427 aagcaacctaattagatcatatagagatgtgtcagaccggttctgtaaagg 4486
    || || || || || || || || || || || || || || || || || || || || ||
Db 9 aacgagattatcatcctcgtcagatgacgaagaacattaccagcgagcgtcgtcg 68
QY 4487 aatccaagaagacatcttcgtgtgagtgctaaagtgtctaccacacagcgtctga 4546
    || || || || || || || || || || || || || || || || || || || || ||
Db 69 actgtaagaacagcctcgtcgtggaagtgctcgtgaggtcgtggaattctcgaactg 128
QY 4547 accaaatgagttatctcaacacactgtttgtcacaagagagcgtgttctcaaat 4606
    || || || || || || || || || || || || || || || || || || || || ||
Db 129 atcagagtaacactactcgtcactcgtcttgcaacactcgtgtctccalcgcgaagt 188
QY 4607 ttccgtgaactctatttgaagaagacagagcagtgctgatttaagtcctcgaagctc 4666
    || || || || || || || || || || || || || || || || || || || || ||
Db 189 ttgagacttactcctcgaagagagtggaacagtgcttgacactatgtccaaagctc 248
QY 4667 tccgacactgtacgtagatcgtgtacataagtcacacacccagtgctcctcttacc 4726
    || || || || || || || || || || || || || || || || || || || || ||
Db 249 tgcaccactcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 308
QY 4727 tactaatgagcaaaccttggagatcgtggagcaactctccagcgttaaatcagtgac 4786
    || || || || || || || || || || || || || || || || || || || || ||
Db 309 tctcatgagaggttcagtttggagccacccagatatttgaagaagtaagaagcagtgaa 368
QY 4787 caatgtactatcctcctcgttgggacacatcagaattttaagaagaattcttaagac 4846
    || || || || || || || || || || || || || || || || || || || || ||
Db 369 ccatgtccctgcgactcttgggtggagagcaccagactttaagaagcaccctgagaa 428
QY 4847 gtctctcaaaagactatattgacatattgcggaagaagatctggaattgaggaagaacact 4906
    || || || || || || || || || || || || || || || || || || || || ||
Db 429 gatccttgaggaacaatttggccattcagaagaagagacacagcagcagatgactcct 488
QY 4907 ttccgtgaagtcacagatcgtgttccaatcctcatatgattcttctgatactgtga 4966
    || || || || || || || || || || || || || || || || || || || || ||
Db 489 ttccaccacagtggtggaagactctctgtatctgatatgacatcatatgacagatga 548
QY 4967 aatgaagaacacacagagagatcctcgaattgtgatgtgaattatgtacagaattgca 5026
```

Db 549 aaatgagggaatttcagagaagatcccgagatcgtatagatctcatgtaacagaattgcca 608
 QY 5027 aaggttaacagacgtcttcaga---gcgattacgtgtgtgaagaacatgaggaagc 5083
 Db 609 agagttacacagacatctccgtctcgtgcgcgtgaactcgtgctccagacatgagagaac 668
 QY 5084 actcagaagaagcaatcactgtgaagctgcacagtgcttaagttccactcagacagacttg 5143
 Db 669 aaccacaagaagaagtgtactacagcgagctgcacatgtgcctgtgtgacgcgcgtgttag 728
 QY 5144 ttgctgaatatttgacatgctcgagagcaggaataatctctctgtggatgtgtaacat 5203
 Db 729 tggctgaatatttgacatgctcgagagcaggaataatctctctgtggatgtgtaacat 788
 QY 5204 ttcaagaatttcatctaatgtttagaagaatctgcggtccagatgatggtgatactc 5263
 Db 789 tccaagaatttcatctaatgtttagaagaatctgcggtccagatgatggtgatactc 848
 QY 5264 cagatgaagaaggtatctgctcgaaataacttacttactgaagacttgggattac 5323
 Db 849 ctgacgagatggtggtgctgcagcgacttacttccagagagtgctgtgttagcctcc 908
 QY 5324 tggacaagcagctgctctctctatggtcgtgcatgtaagcagttaatgaagttt 5383
 Db 909 tggagcagcgccgagctcttcagcacgaggttatatagaagttaatgagctc 968
 QY 5384 acaaaagtacttacttacttacttacttacttacttacttacttacttacttacttactt 5443
 Db 969 acaaaagtacttacttacttacttacttacttacttacttacttacttacttactt 1028
 QY 5444 atgttaaatctcaagaagcatctcaaaaatgttcatcagaagactcgtcgagagcga 5503
 Db 1029 acagcaagctgtcagagagcgtctcgacagcatcgttaacaagatcat-----aaga 1082
 QY 5504 tgtttggaacctatttctgtgtgtttttagaacaagaatctcgaggatttggatgac 5563
 Db 1083 tgtttggaacctatttctgtgtgtttttagaacaagaatctcgaggatttggatgac 1142
 QY 5564 aagaatttgttacaagaagcgtcgtaataacaaactcgagagatctacagatttg 5623
 Db 1143 aggaagtctgtctacaagaagcgtcgtaataacaaactctcctcgtacatctcag 1202
 QY 5624 aggaatttcaagaagaagatttgaagagatgtgttgaagtaacaagacttaac 5683
 Db 1203 aggcatttaigtcaagtgtttgtgcagaattgttgaagtgatgaataacatccactc 1262
 QY 5684 ctgtagacaagtgttaaatagatcctaacaagcatatattcagattatgtgagc 5743
 Db 1263 ctgtagacaagaagcaggtgtgacttaacaagcctacatagatcttgtgagc 1322
 QY 5744 catacttggacacatattagatgaagagacatatactatttgcacaaaattacac 5803
 Db 1323 cctacttggatgattagatgaagagacaggtgtacactacttgaagaatttcaac 1382
 QY 5804 ttgcgtgattcatctgactgtacaccccttaacttagatgagcgcgtcccatgggaactc 5863
 Db 1383 tccggaggttcatctgactgtacaccccttcaaccttgaagggcggtcgtgggagcgtgc 1442
 QY 5864 atgaacaatcaaaaggaagacacttgaactagctcactgaagcttccctatataaa 5923
 Db 1443 atgaacagctacagaagaagacagctcctgacacactatgacagcctccctacataga 1502
 QY 5924 caagggtcaattctcatataaagaagatcatcttaacccaattgaattgtctatg 5983
 Db 1503 ccagatcagcgtctcaccagaagaagagtggttgttgcacccagattgaattgcatcatt 1562
 QY 5984 aggaacatgcagaaaaagacacagagctgtgcatttgcacaacatcagatcccgagacc 6043
 Db 1563 aagaaatgaagaagaagaccctgcagcttagcagttgcataacacagagagcgcctgattg 1622
 QY 6044 ccaaaatgcttcatagatgtaactcagggatctgttaggcacacagttgaaggagcctt 6103

Db 1623 caaagatcttcagatggtgtgcgaagcctctgtggagctactgttaaatcaggagccac 1682
 QY 6104 tggaaagtgtccagagttttctgtctgaataactaagtgacccaagcttttaagacac 6163
 Db 1683 tggaaagtgtccagagttttctgtctgaataactaagtgacccaagcttttaagacac 1742
 QY 6164 ataaataactgcagctcgtctgtttaaagatttcaaaaagttggaagctttaaaga 6223
 Db 1743 acaacaagttgaggtatgtctttaaagatttcaaaaagttggaagctttaaaga 1802
 QY 6224 aaataagagcttaattggccggttcaaaaagttggaagctttaaaga--aact 6281
 Db 1803 aaacaagcgtctcatcagcagagagagatatacagcaggaactcaaaaagaact 1862
 QY 6282 atctgcgttaagaagagccctacagccctagatc--acagaagtcctcagttatcca 6339
 Db 1863 atacaagcttaaaagagaactcagccaatgatcagcgaagaatctcagaactgtaaa 1922
 QY 6340 agcagatct 6349
 Db 1923 agccaatatt 1932

RESULT 11

ABLI7736/c
ID ABLI7736 standard; DNA: 9389 BP.

AC ABLI7736;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 4681.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ds.

XX Drosophila melanogaster.

OS WO200171042-A2.

PN 23-MAR-2001; 2001WO-US09231.

PF 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PMD, Myers EM;

XX WPI; 2001-656860/75.

DR New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions -

PS Claim 1; SEQ ID NO 4681; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (ABLI16176-ABLI30511), expressed DNA

XX sequences (ABLI1840-ABLI16175) and the encoded proteins

XX (ABBI7737-ABBI72072).

XX The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 9389 BP; 2710 A; 1842 C; 1849 G; 2988 T; 0 other;

Query Match 9.1%; Score 621.6; DB 23; Length 9389;
Best Local Similarity 49.2%; Pred. No. 8.9e-146;
Matches 3017; Conservative 0; Mismatches 2424; Indels 691; Gaps 25;

QY 536 cctgtagatattgacttgaaaattcaactcctgatatgcttgcttcccaattactg 655
DB 7494 CTTGGGCTAGCTTGATGATTCGAGTTGCTGTTAATGACCCACTAATGCCCAATTTGGCTG 7435
QY 656 atcgacatcccaatgaagaatagacgctcgaatgatgaccagaagaatacaacgctc 715
DB 7434 ACAATGTACCTCCGAGCATATGTATCAATCCAAATGAAGACGCCGGAGCAAAATCGAC 7375
QY 716 acaagaactcttgcttgatcatcatcaccagaatgaagaacaacataagagctta 775
DB 7374 AAATGGCTTTTTCACATGTATCCGGAAATCGAAGCGAGAAACATTTGAAGCGCCGC 7315
QY 776 gtgttcctgatataccagaacacatttgtaagaagactctgttaaatgcttatac 835
DB 7314 TGCTCGCAGCATTCACATTCATGACATATGGCTCATCGCATCAAGTTAACTGCCTACAGC 7255
QY 836 tcaagttgaaattgaattgaaccattttgcaagtttgcttatgatgtaag 895
DB 7234 TTCCGCTTAGCTGGAATGAGCCAAATTTTGATCGATGGCTATTTATGATGCTTAAGG 7195
QY 896 aaaaagaagaattcagaagaactttaatttgaccttaattcgaagaatgaagaagt 955
DB 7194 AACGCAAAAATTTCTGAATAATTTCTATTTTGACATGAAGCAGATATGTTAAAGCA 7135
QY 936 tgttaagtcacatgataccacctgctgcacattaccacctggaagaatcagaactttt 1015
DB 7134 TGTGTCTCAAGTATGTACAGTGCAGCCGACATAGTACCAAGCAGCACCTACATATTCG 7075
QY 1016 ctatcactatccctcccaagatgtttctgttaataagctgaagaatgtaagc 1075
DB 7074 AAATTTCAATCCAGTAAACGATTTATTTCTGCGATTCGCTGGAATAAATTTTGACG- 7014
QY 1076 aagagacattggaagagtgtcagaaccatalatgatttcaaaagacagatgcca 1135
DB 7015 -GGTACATTAACAAATTCGCGAGCGCTATCTCAAGAAAGACAAAG----- 6988
QY 1136 agaataaagaactgtagaagaactgaagaatcagaagatcagltttgccaagaactg 1195
DB 6969 -----ATTAATATCCGAGAAAGTAATCAACGCTGTGACTACTGTAACAGCTAG 6916
QY 1196 ggaatatcgatcgcttcttgcttgactgcaatccatttaataatgtt----- 1249
DB 6915 GCMAATACCGAATGCCGTTTGATGACAGAGAAATCTACCTAACAAATGTTTAAATGCGC 6856
QY 1230 ----- 1249
DB 6855 ATTAATTTGAAAGCAAAAGCGTGGCGCGGGAACGGGATTCCTTTGGAATGCTTGTG 6796
QY 1230 -----gcaagtctggaggtttggaagaagatctcagaagtagaatacagta 1237
DB 6795 GCAGACAGCTTGGGCGACACGCCAATCTCAATAGTTAGATGAAAATCTTCCCAAGTA 6736
QY 1298 ctgsgaagcgaagaaggtctgtgtcagagagaggaattcagatgtttgagagagat 1357
DB 6735 GCTTTGATCACTAAGCGGAAAAGCTAAAGATATGAGCGGAACCTTACAGAGAGTGTT 6676
QY 1358 caactgaagaagcaaca-----gtgagatgatctgttaactga 1399
DB 6675 CCCTTGAACGGAAGAAAAGACCGATCTGTCGCCAGATGATTTTCCAAACGTTGAG 6616
QY 1400 cgaagttcgcacagctctctcagatgacaattttttaagcagaagaagacgct 1459
DB 6615 AAACTTTGCTCTATCACTTAACATACAGCTTCTTTAAACAAAGAGCAGATATAA 6556
QY 1460 taagatgaagaatctcacaactcctgtgtatgtagaagaagcactctctctac 1519
DB 6555 TGAAGAGAGAGATTTGTACAAATTTTACAGAGATTAAACGACCGAGCTCAGTTATGA 6496

QY 1520 ggcgactaagaccattatacagctcagctcaagataagacatttctccgcgacctgaatlc 1579
DB 6495 AGAAGTAAACAATGCATACCTGGTTCTATTAAATATGAATATACACTCTGCGAGGAAG 6436
QY 1580 cccattatgctcaactccgagctgcttcaagtgaagcttaacctgaagtagagta 1639
DB 6435 CCAAGATATGATTAAAGCAGAGATAGCACTATAAACCTCAAAATGCAAGATATATGTC 6376
QY 1640 gacctcagagaatcttagagttcccgcaaggagtgattatgttccaaacactact 1699
DB 6375 GCCCAGTTAAGAAATCTTATGAAATTTCCGATCGCGATTTTCAATCCATTTATAGCT 6316
QY 1700 acagaatctctcctcatatataccctcagaagcttcaatttgcacalcgtaaggttctg 1759
DB 6315 ATGCAATCTCTATTTGTTGTTGTCACCAAGAGAGTAAATTTCTTGCGTGCAGAGTTG 6256
QY 1760 ctagaataataacagtgaaagtlccagttatgtatgtagagagatccaagaatgcagc 1819
DB 6255 CTGGAACATTTGCCGTTCCGCTCAACTAATGCTGGAAGAACGCCAAAGATGCGGTCA 6196
QY 1820 cgttaactcttgtaactcagctgtgtcagaatcttcaaggaagcctatagacgctag 1879
DB 6195 ATGCAATCTACGCAAGTCTCTTGTCCCAATTTTCTCTGAGATTTTACAGCTGTCA 6136
QY 1880 tatacataacaggtctcctgatttcaatgaagaatccaaggtlaagcttccgtact 1939
DB 6135 ATTATCTAATCAATGCCATCATCTTACGATGAATTAATAATTTGCGCTGCATCA 6076
QY 1940 taactgacatcaactcagcttcttacttataatgtatgtatgttcaaaaaa 1999
DB 6075 TAAAGCAGATCAACCATTTATTTGTCACATTTACATGTTTCATCTCAAAAGAACCC 6016
QY 2000 at-----actcctctgaaacacagtgtaataatcagatgatacagctcaga 2050
DB 6015 AGCAATTTGCAACCGCTAGTAAAGATCTCTATGCGCTACACATGTTGCCCTTACTGGAAG 5956
QY 2051 atggaagcttgaagactgagcagtttgccttgctgagcttccatctgaaacacacacag 2110
DB 5955 ATGGAATACTTAAGTTGGGAATTTAATCTGCTGTTATGGAATACCGCGCGA 5896
QY 2111 cttaactctgactgctcctcaggttccctcactcagctgagaaatggtatagataatcaca 2170
DB 5895 ATTACTATTTATACCAACCAAGTTACCTGCTGGAATTAATGCGTGCACAAATCA 5836
QY 2171 aaggtgtttaaagt 2230
DB 5835 GAGCGGTTTCTATTATATGTAAGACAGTCAACCGCAATTCATTTGATTTGATTTCTTCC 5776
QY 2231 ttgacaatcttctgctcgt 2290
DB 5775 TCGATGATTTCTTAAATTTGGGAATATCTAGACACTGGAATATATCTTCCATATTTG 5716
QY 2291 gggacatggaatcagtgaaataacttagaataatgaatgaagagcagatattcagac 2350
DB 5715 G-----CGAATAATATATGAAACAAATTAAGAAATCTGCTGTGATA 5671
QY 2351 tgaattatccagcagtgaaacaggtgttcagatcttctatctctgctagataaactga 2410
DB 5670 TCGAATATGCAATTCGAGACCATTTAGTAGAGCATCTTACCATTTAGTCTTGGATTA 5611
QY 2411 tactttagttatagactcctcgtgcatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2470
DB 5610 TCGAATTTGTTAGTTGTCACCAAGTTGTTGAGCAAGCAAGTGTCTGGATCTACAG 5551
QY 2471 ctlttgagccatgagcatcaataataacagcttcaaaaaacttggaagaatcatgt 2530
DB 5550 TTTTCGAGTCTTCTTGTGTTGTTATCTATCTTTATCAATACTTAAACGAGATCAATAG 5491
QY 2531 accagcagtgagcaaaagcctctgcatcataatcaatcaatgttctccgctcaaa 2590
DB 5490 ACCAATACGCTGCGCAAGTTGCTATCAACATATGTGCACTTTCAATGTAATATTCGCC 5431
QY 2591 ataactaccataatcatacaccaggtcctgtggtgtgtgtgtgtgtgtgtgtgtgtgtgt 2650

Db 5430 A----- 5428
Qy 2651 ccacaaatgctagatctgcgtgagacccgtaacgcttaattaaatcgttctcgaaac 2710
Db 5429 -----CC 5428
Qy 2711 ttagtaataagcaatccagatatatctggagctcccaatccacagatgatgaagtcgat 2770
Db 5427 CCTTCAAGCAAAAGCAGCTAATCATGAGTCGACAGCAACAGACAA----- 5377
Qy 2771 caatacggaggaaggttagatcgctccaatccctgggttaacaacggtggtccaa 2830
Db 5378 -----CTTGCATTTAAGCAATCTGACACCATATATGATATGTTTATGCAAGTGCCGCA 5324
Qy 2831 aagctgccccatgggagatcccaaccccaatcccaatgagcaagatcaacaaggtatgagtc 2890
Db 5323 AGCTTGGATCG-----CAAGGTAAATATCATATTTGTTAAAGCTTTAAATA 5276
Qy 2891 gaagtgttaatcgtatgtcttcgcacacagagagctcaagttcttacaacaatlaacg 2950
Db 5275 TAACTCTTTTCTTTCAGAGCTGTCCATAGA-----TCTCTTCACTCGATGCGCTG 5225
Qy 2951 gaagcttacaactaaagacttccagagagagctggcttgcaagtggttgca 3010
Db 5224 CACGGGATGTCACAGTGGCGCTTCCAGAGAAATTAAGCTTGATGGGTGTGCAA 5165
Qy 3011 gtagagagcttcggaatcagcttgcacaagaagctggtcttcttgaataatgtaa 3070
Db 5164 GGGAAAGCAGCTGACTAGCCATGTCAAATTCGTGTTTATTTAGCTTATTTGCA 5105
Qy 3071 agagatggtgcacattatattcttaataatgagcttcgaagctccaagaaagctgt 3130
Db 5104 AATCATATGATGTAACATTTGATGTTCATAATCTTTAAATGTCGCCGCAACATCGAT 5045
Qy 3131 tcccaagacgttcatgtagatgacatgcagctctgtcagaacgagtgtgtagatag 3190
Db 5044 TTCCACATAGTTATGATGATCATCCACCTGTCACATTAATTAACAAAGAGTG 4985
Qy 3191 ttccaagattcagaagacagaaatggtgagagacatcaatcaagcttgcattct 3250
Db 4984 TCGGTACCAACAGCAACGACCTAAATTTAGCCAGCTCATTTAAATGCAAGTTTAAAGCTTT 4925
Qy 3251 ttctcaatgatactgtgtctgtatgcagagagatgttcttgaagcttaaaagctct 3310
Db 4924 TTATATTTGACATACTAAGCATCATGATGAGTGGGATTTGTTTGGCCTCATCAAAAGCT 4865
Qy 3311 gctataaacaagtgctccaagaagcttcaatcaatccgaatccagtggtctgtctct 3370
Db 4864 ACACAAAGGTTCTGATTTCTAATAAATGCTCATATACCG-----GATCTAATGAACT 4814
Qy 3371 tgaagctgattttcttaagacatcatcgaatcgaatgagacatgattcatatcaatcac 3430
Db 4813 ATTAATAATGACTTTTGTGAGAAATTTGTGTAGTCATGACATTTTGTAGCCTTAAACTTGC 4754
Qy 3431 cctgcagcttacttactccacatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 3490
Db 4753 CGTTTGTGTAATCATACAAATGTAACAGGCGCATGACGCCACGCAAGCAACAAAGCT 4694
Qy 3491 agagttcgtgatttctcagaatgtaacaagcaaaagattgcaataatgttgaatat 3550
Db 4693 CGAGGACAGCCAACTTCTGCGTAAGTAATAAATAACATACACTTTAATTTTACTTAT 4634
Qy 3551 cgtgccttccgcgaacagcatattgtgcagga----- 3585
Db 4633 TTGTGGCTTTCTTTAAAGGATCACTAGAGAGAGCCCTTACAGCCGCACTTACTCAAGA 4574
Qy 3586 -----ctgtgttlaacagagctggtctgcatltttagaccctgaltgc 3626
Db 4573 ATTCTCCAGCAACAATTTTCTAGTGCAGCTTGTCTGAGCAGACTTAGACAGAAATAGGA 4514
Qy 3627 tgaagagctgttctgattgcatagaagaatcatcaatagttacaaattactctcag 3686

Db 4513 GGTCCAAACCCCATTTACATGGAAGAAAGCATTAAGGTGTATTTCCGAACCTGATGACATC 4454
Qy 3687 tcaagactcagagcccggtactctgaccctcgatataaagcccgagtggttga 3746
Db 4453 TCACGATTTGGACCGGGATATAGCAACTGATGCAGTCAAGAGGTGGCTCTTTATA 4394
Qy 3747 tctaccctgattgatttctatcatalggaactgtaactcaagctgatatgatttacaagaac 3806
Db 4393 TATACCGCTACTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4334
Qy 3807 tcaacatcaacggaagaccatattgtatagccaactgatatgatatgaagtgagagcg 3866
Db 4333 GCAGGATCGTCTGAGCAAAATCGGTCACT-----TGAAGATCATCAAGGCGCACATCA 4280
Qy 3867 aagatgataagccagacggttgcacatggaatcgaagaggaactggtcccttaactaac 3926
Db 4279 AACTATTACAGCATCAACCAATTAATCCGGAAGTGCA-----TTTGCATATCGG 4230
Qy 3927 aagcctgagagttccctcctcaagctcaagagtgagcaacacacatccttccagc 3986
Db 4229 GAATCGTCCCTACTCTCTACTTGAATGACAACTTAAACAAATTTGCCCTCAGCTCGG 4170
Qy 3987 agaatcagatcgaagccttltgaltcgttacttgggttccaaaatgcaatgaatgaac 4046
Db 4169 AGAAT--ACCGGCATTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 4112
Qy 4047 agttcagaagatggttccatcgtctcagcttgcagctcaaccgctatgattct 4106
Db 4111 CGTTTATATCGCTGCTGCTTATGATTTAAGTCCGCATATAGTCCACCAATGCTTCAAG 4052
Qy 4107 gcttatactcgt 4166
Db 4051 TGTAACTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 4007
Qy 4167 tagcttgaacttgaagaatcaaaagacatgagagcaagcttgaagaagctatcttgc 4226
Db 4006 CGTGGCCACCTTAAAGCAAGATACACAAATTTTATAGCAAACTGATCTACTGATGT 3947
Qy 4227 gagcatagtggtccagcaagaatgltgtagcggaagccgagagcaactcgaagaagccc 4286
Db 3946 TAAAGAAAGCTGGAAGAAATGTTATGAGGACACAAATTCAGCAGTTATGATTTAATAA 3887
Qy 4287 atctgagagcttcttgaagaatcaagaatlttgaagtggaagaatgatactgactg 4346
Db 3886 TCGTCAAAAGATGCTAATTTCAACAGAAATTTCCGATGAGGAAAGACCAAGATGCAAT 3827
Qy 4347 gcg-----tcaaaacacagagaagcttgaacaaatcaagaagcagagaattgaacagagcact 4403
Db 3826 TCGCTCTCAATATGCGGATGCTGTGGAAAGAAAGTGAAGCATGAACTGAATTAAGTCATTT 3767
Qy 4404 gattgatggaacatcgtctcagaagaacaaactcaatcatctttagatattagatgattg 4463
Db 3766 CATCGAAGGCTCTTTGGCAACTGAAGTTCACATTTGATGATTAACCTTGAATAATAT 3707
Qy 4464 tgttcagagccgttcttgaacggaatccaaaggaagcattcttgcgtgagtgctaaagc 4523
Db 3706 TGTTC-----TGTGCTGCCAACCTTTATCAACATCTTCTTGGAAGCTGTGTAAGAT 3653
Qy 4524 gctactacacagcatgctgtataccaagtgagttatctacaaactggttgcctac 4583
Db 3652 GCTTCTTCATCTTGTGCCGGAATCATCTGTTTGGCCCTGCCAAACTGTTTGGCCCTC 3593
Qy 4584 acagagagccttgggttcaagttctcgaactcaatctttagaagaagacagagagtg 4643
Db 3592 CCAAGCGGCTTATCTTTAATTTCCAAACTTATTTATTTGATGACGACACCGATATATG 3533
Qy 4644 tgcgtatltatgcctcagagcttccgacacatgtaacagtagatcgtgtacacagtc 4703
Db 3532 TGGGATTTGTGCTATATTTACTGAACATTTGTGGTCCCATATGCCAGAAATGACATC 3473
Qy 4704 acaaccagatgctcccttactactaatgagcaaaacttggatgggaat----- 4758
Db 3472 ACAGGAGCTGCATCACTATATTTACTAATGAGACAAATAATTTGAATAATGAAATGTAAG 3413

QY 4759 -----aact 4762
Db 3412 TTTTATTTATGTTAGTTTATTTAACTAATTAATGGTCACTTATTCGTATAGAACT 1111
QY 4763 ttgcagaggttaaaatgcaggtacaaatgcatcactccctctgtyggcacatctcaga 4822
Db 3352 TTGGCCAGTTTAAGATGCAAGTGCAGATGCTTTAAAGCTCCCTAGTTGGAAGAATGTCG 3293
QY 4823 atttaataagaattctcttaagacgttctctaaagactataatgacataatgcagaaga 4882
Db 3392 TCTTTATGAGCAATCCTTACGGCTGGCTTAAACCGTCCTGTTTATGCGAAACCG 3233
QY 4883 atccgaattgaagaaacaacattctctgatacaggtccaggtctgtyttcaatctc 4942
Db 3232 ACCTCGACCTGACGAGTACTCTTTCTCTGACACAGTACAGATTTCTTTTAATCTGC 3173
QY 4943 atatgattctctctgatactgtygaanaatgaagaaacacagagagatcctgaaatgtga 5002
Db 3172 ATATGATCTCTGATGATCTGTTAAATGAAGAGTATCGGAAGACCCAGAAATGTGC 3113
QY 5003 ttgatacaatgataagattggccaaggtttaccagacctctccaga---gcgattgacct 5059
Db 3112 TTGACCTCATGTAATCGTATTCGCAAGGATACCAAAATTAACCTGATCTAGACTGCTT 3053
QY 5060 gttgcagaacatgycagcagcaagcactcagacagcaatcattgctgaagctgacagct 5119
Db 3052 GGTGGAATAATATGGCTAATAAAACACCGGACGACGCAAAATCACACGGAAGCAGCCATGT 2993
QY 5120 gctagtcactcagcagcagctgtygctgaataattgaagatgctgagagccagaat 5179
Db 2992 GTTATGTACATGCTGCTTCTTTAGTTCTGAAATCTTACATGTTGTCACAAATTA 2933
QY 5180 attctcgtgaggtgtygtaacattcagaatattcatctaaatgcttgaagaatctg 5239
Db 2932 ATTTGCCCTGTGAGCTGTAAGTTTCAACCAATTTCTCCAAACACTATATGAGAGTCGG 2873
QY 5240 cagctcagatgattgctatccagatgagaagaatgctgctcgcgaataactta 5299
Db 2872 CCGATCGGATGATGCTGTAGTCCCGGCAAGATGATGCTGCTAGGAATATTTCA 2813
QY 5300 ctgagtcagagctgtyggattactgtaaacagaagcagctgtctctctctatgctgga 5359
Db 2812 CTGAACACTGGGTTAAGGCCCTTCTGGAAGAACCTCCAAATTTCTTCAAGTTGCTGCA 2753
QY 5360 tctatgaagcaggttaatgaatttacaagaagctattcctcattcagaagcttaatcg 5419
Db 2752 TGTATGAAGCAATGACAGAGTGTACAAATTTTAATACCATATGCGAGGCTAACAGAG 2693
QY 5420 atgcaaaagaactatccacaattcattgtaactcaagaagcatt----- 5465
Db 2692 ATTTTCAAAAGCTAAGCAAAATTCATGSCAAATTCAGAGAGCTTTAATGCAATATCCC 2633
QY 5466 -----cagcaaaattgtcatcagagact 5490
Db 2632 AACTACAGGTAAACATATTTGTATAATTTTACCAACGGAATAATATATTTATATAA 2573
QY 5491 ggcctggagggagatgtyggacatttgcgtgtygttttgaacaaagctcg 5550
Db 2572 CAGGTTAAGAGATTTTGTGAACTACTTGTGTGGCTTCTATGCGGGAATTTGGG 2513
QY 5551 gatttgaatgaacaagaatttcttatacagaagagcctcaataacaactgycagagata 5610
Db 2512 GACTTGGATGACGAGGATTCATTTAATAAGACCAACATTTGAGAGAGTCCCGGAATA 2453
QY 5611 tct----- 5613
Db 2452 TTTTACTCGCTTCAAGTATATATTCGAAATTTGAAATTAAGACTATCAATTTTGT 2393
QY 5614 -cacagattgagggatttaccgagaagaatttgagagagatgtygttgaagtaacaa 5672
Db 2392 CAACATACGTTAAGAACTTTACACTGAAAGATGTGGACCGGACTCTGTGATATCATTTAA 2333

QY 5673 agactctaacctgtagaagaatgtaaatagatccatacaagaagcatalattcagatcac 5732
Db 2332 AGATTCGATFACCGTTGATATTAATAGCTGGATCCGATTAAGGCTTACATTCAAATTAC 2273
QY 5733 ctatgtagacatcatttgcacatatgagatgaagaagaatcactatctgcagaa 5792
Db 2272 TTATGTTGAACCTTACTTTGAAACATATGAATVCGTCATGCTGAGACATCTTTGAGCG 2213
QY 5793 aaattacacatctcgt----- 5808
Db 2212 GAATTTCAATATATGATGATGTGAATTAACAGATTAATTTAATCGAAATTTAAATTA 2153
QY 5809 -----cgattatgtaactgtaacaccttacttag 5839
Db 2152 TATTTTAACTTTTGTATTATGCAAGACGTTTATATATGCGCCGCTTTTACTATAA 2093
QY 5840 atggccgtgcccattggaggaattcagaacatctcaaaagaagaacatctgtaactcgt 5899
Db 2092 ATGGAAGGCGACATGGCGAATTAATGAGCACTGTAAACGGAACGATTTTGAACACGG 2033
QY 5900 ctcatgctctctctatataaacaaggtcaatgtaactcatalaagaagaagatcacct 5959
Db 2032 CGATATCTTTCCATACGTAAAACTCCGATTAATGTAATTTAGCCGACGCAATATGTC 1973
QY 5960 taacacaaatgtaagtgctatg----- 5983
Db 1972 TGGAAACGATTTGAAGTGAAGTGAATTTGAAGTATGTTGAATCCCAAAAGTTAATA 1913
QY 5984 -----agacatgcaagaagaacacagagattgcaattgcaaac 6025
Db 1912 AATTAACGTTATTTGCTTGGATATTTCAAAAAAACAATTTGGAATTTGCGAGCTGCCAGA 1853
QY 6026 atcagatcccgacagaccccaaaatgctcagatgtaactcagagatctgtagcacca 6085
Db 1852 ATCAAGACCCAGCAGACCCCAAAATTTTCAATATGTAAGTGAAGGATATTTGAACCA 1793
QY 6086 cagtgatacagggccttggagatgcccaggttttctgctgaataacacatgtagcc 6145
Db 1792 CTGTATACCAAGACCGATGAAGTGAATGCGATGCTCTTCCAAATTTATTCGACGGA 1733
QY 6146 caa---agctctcagaacatcataaactcgcagctcgtcttaagaatttactaaa 6202
Db 1732 CAACCTGTACCGCAAAACACCAAAACAACTTCGGTGTGCTTTGGCGGTTTCAAGC 1673
QY 6203 ggtgtaagatgcttgaagaanaaataagagcttaattgycgcggttcaaaagatgac 6262
Db 1672 GTTGTGCTGATGCTTTGAAAAAGATCGCATCTAATCTTCAATCAAAAAAGATTACC 1613
QY 6263 aaagggaattg 6274
Db 1612 AACGAGAACTGG 1601

RESULT 12
AAH99572
ID AAH99572 standard; cDNA; 1652 BP.
XX
XX AAH99572;
AC
XX
XX 16-OCT-2001 (first entry)
DT
XX
XX Human protein encoding cDNA sequence SEQ ID NO:407.
DE
XX Human: cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiac; central nervous system; virucide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; aneasmic; anaemia;
KW antiagregant; haemostatic; vulnery; antilicer; osteopathic; eczema;
KW dermatological; allergic; antislumatic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;

SEQUENCE	13
ID	AAI61331 standard; cDNA: 2427 BP.
XX	
AC	AAI61331;
DE	22-OCT-2001 (first entry)
XX	
DE	Human polynucleotide SEQ ID NO 5320.
XX	
KM	Human; neotropic; immunosuppressant; cytoskeletal; gene therapy; cancer;
KM	peripheral nervous system; neuropathy; central nervous system; CNS;
KM	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KM	altheropathic lateral sclerosis; Shy-Dragger Syndrome; chemotactic;
KM	chemokine; thrombotic; drug screening; arthritis; inflammation;
KM	leukemia; ss.
OS	Homo sapiens.
XX	
PN	WO20015312-A1.
XX	
PD	26-JUL-2001.
PF	26-DEC-2000; 2000WO-US34263.
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-0598042.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0727344.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
P1	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
P1	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
P1	Zhao QA, Zhou P, Goodrich R, Dimanac RT;
XX	
DR	WPI: 2001-442253/47.
DR	P-PSTB; AAM42175.
XX	
PT	Novel nucleic acids and polypeptides, useful for treating disorders
PT	such as central nervous system injuries -
PS	Claim 1; SEQ ID NO 5320; 10078pp; English.
XX	
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC	the encoded polypeptides (AAM38642-AAM42213) with neotropic,
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	localised neuropathies and central nervous system diseases, such as
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	utilisation of the activities such as: immune system suppression,
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC	and thrombotic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukemias and
CC	C.N.S. disorders.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification.
XX	
SEQUENCE	2427 BP; 716 A; 508 C; 524 G; 679 T; 0 other;

Query Match	6.7%;	Score 459.8;	DB 22;	Length 2427;
Best Local Similarity	68.8%;	Pred. No. 2.8e-105;		
Matches 678; Conservative	0;	Mismatches 297;	Indels 10;	Gaps 3;

QY	5369	cagttaaatgaagtttcaaaagttactctatctccattccattgaagctaatcgtgatgtgaaga	5428
DB	4	caatttaattgaggtgttcaaaagctgtgtccatccccaatccctaaagtcgcatcgtgaattccgga	63
QY	5429	aactatccaaactatcagtggaactcttcaagaagcattccagcgaanaattgttcaatcagagta	5468
DB	64	agctgcagatctcatccacaaagaagctgcagagagccttcgacagatctgttaacaagatc	123
QY	5489	ctggctcggagcgcgaatgtttgcacotattttccgtgtgtgttttaagaaaccaagtctg	5548
DB	124	ata-----agagatgttttggaaccttaccgagtttggttcttcttgatccaaatttg	177
QY	5549	ggagtttgtaggaacaaagaatttggtttacaaaggagcgttcataacaaactatgtgaaga	5608
DB	178	gggattcttgtagaacaaagaatttgcctcacaaagacgtccgaattccacaagcttccctgaga	237
QY	5609	tatctcacagatttggaaggatttctacggagaaagatttggaaggatgtgtgtgaagaa	5668
DB	238	tctcacatagagcttagagcattttatgtgtcaatgttttgggtgcagaatttgggaagtga	297
QY	5669	tcaaagacttaatccctgtgtgacaaagtgttaaatatagataccatacaagcatabatcaga	5728
DB	298	ttaaagactccacatccctgtgacaaaaccaaagtgtgattcttcaacaaggcttacaatacaga	357
QY	5729	ttacctatgtgagccatctactttgacacatatgattgtagaagaacgaatcacattcttg	5788
DB	358	tcaattctgtgagccctactcttctgtgattgattagattagattgaagaacggttcacatatcttg	417
QY	5789	acaaaataatacatcttcctgcgtacattcactgtactgtacaccccttactttagatgtgcgtg	5848
DB	418	agaagaatttcaacctccctgcggaggtttcactgtacacaccccggtttcacccctggaggcgcg	477
QY	5849	cccatgggggaacttcatgacaacaattcaaaaaggagccattctgactacgtctcagct	5908
DB	478	ctcgggggagctcgtcatgagcaggtacagaagaacacagctccctgacacattatgcagcct	537
QY	5909	ttccttatattaaacaagaaggctcatctatccataaagaagatcatcttaaaccaa	5968
DB	538	tcacctaatatcaagaaccaagatcaatcgctcatccagaaggagagattgttttgacaacga	597
QY	5969	ttgaaagtgtctattgagagacatgcagaaaaagacacagaagattgtgcatttgacacatc	6028
DB	598	ttgaagttgtccattggaagacatgagagaagaagccctgcagtttagattgtgcattaac	657
QY	6029	aggtatcccgacgaccccaaaatgctctcagttgtactccaaggatcttgaagcaccacaag	6088
DB	658	aggagccgcgtgattgacaaagaatgcttcaagttgtcgtcgaaggctctctgtggagactcag	717
QY	6089	tgaattcaagggtccttttgaaagtgcacagttttctcgtctgaaataaccttagtgaccaa	6148
DB	718	taaatccagggtccactgtgaagtgcacagatgttttcttgctgaatactcctgtcatccaa	777
QY	6149	agctcttcagaacatcataataaactcgcactcgtctgtttaaagaatttactaaagggtgtg	6208
DB	778	aactctatcagcatcacacaacaagtttggtgtttagtcttaagatcatcatcagatgtg	837
QY	6209	aagatgctttagaanaaaataagaagcttaattgtggccggtttcaaaaagggtgttcaaaagg	6266
DB	838	gtgaagcgtgtgagaaaaaacaagcgtctctacacggccagacccagagggaatactcagcaag	897
QY	6269	aattgggga--aactatcttgccttaaaaggccctacagccctagatc--acagaag	6324
DB	898	aactcaaaaagaactatatacaagttaaagaagaaactcctgccaatgatctcgggcggaaa	957
QY	6325	tcctcagttatccagccagttat	6349
DB	958	ttccagaactgttacaagccaattat	982

RESULT	14
AAI59545	
ID	AAI59545 standard; cDNA; 2433 BP.

XX	AA159545;
AC	
AD	22-OCT-2001 (first entry)
DE	Human polynucleotide SEQ ID NO 1748.
XX	
KW	Human; nootropic; immunosuppressant; cyostatic; gene therapy; cancer;
KW	peripheral nervous system; neuropathy; central nervous system; CNS;
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX	leukaemia; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO20015312-A1.
XX	
PD	26-JUL-2001.
XX	
PF	26-DEC-2000; 2000MO-US34263.
XX	
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-0598042.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0727344.
XX	
PA	(HYSEQ-) HYSEQ INC.
XX	
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX	
DR	WPI; 2001-442253/47.
XX	
PT	P-PSDB; AAM40389.
XX	
PT	Novel nucleic acids and polypeptides, useful for treating disorders
PT	such as central nervous system injuries -
PS	Claim 1; SEQ ID NO 1748; 10078pp; English.
XX	
XX	The invention relates to human nucleic acids (AA157798-AA161369) and
CC	the encoded polypeptides (AAM38642-AAAM42213) with nootropic.
CC	Immunosuppressant and cyostatic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	utilisation of the activities such as: Immune system suppression,
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemias and
CC	C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification.
XX	
XQ	Sequence 2433 BP; 714 A; 505 G; 528 G; 686 T; 0 other;

Query Match	6.4%;	Score 434.4;	DB 22;	Length 2433;
Best Local Similarity	70.8%;	Pred. No. 7e-99;		
Matches 606;	Conservative 0;	Mismatches 246;	Indels 4;	Gaps 2

Oy 5498 agcgcgatttggcacctaattcgltgttgatttatgacccaagtctggagattcgg 5557
|| ||||||| |||| | || ||||||| |||| |||| | |||||||
Db 139 agagaatvtttggaacctacttcgaagtgttctcttgatccaaatttggagattcgg 198

QY	5558	atgacaagaatttggtttacaagaagcgtccataacccaacttcgagaagatattaca	5617
Db	199	atgaacagggaatttgccttacaagaagcctgcaattaccaagcttccttgagattctaca	258
QY	5618	gattggaaggatctttaaagggaagaagatttggaagagatctggtttgaagatccaagact	5677
Db	259	gacttagaagcattttagtgcataatgltttggtgcagaatttctggaagatgataaagact	318
QY	5678	ctaatccgttagacaagtgtaaatctagatccctacaagaagcatatattcagattaccta	5737
Db	319	ccactccgttgacaacaacaagcttgatctcctaacaagaagcctacatacagatcaactctg	378
QY	5738	tggagccataacttttgacaactatagatagtagaagacagaatctacactattctgcacaaat	5797
Db	379	tggagccctactcttgatgtagtatgtagtgaagaagcagggttcacataactcttgagaagaatt	438
QY	5798	acaactcttcgcgtatctatgtactgtacacaccttctctttagatggccctgaccctgag	5857
Db	439	tcaactctcggaggttcaatgttacaacccccgttcaaccttggaagggtgcgctcgggag	498
QY	5858	aactcttgaaacaattcaaaaagaagaccattctgactagctatgctcatgtcttcccttata	5917
Db	499	agctgcattgagcagtagacagaagaacacagctctgaccactatgcaagcgtccctccata	558
QY	5918	ttaaacaagaggttcaatgtctactccctacaagaagatcatcttcaacaaccaattgaattg	5977
Db	559	tcaagaccacagatctatagcgtatccagaagagagagattgttcttgacaaccgattgaagttg	618
QY	5978	ctatttgaggaatgcagaanaagacacagagattgctgatttgcaacacatcagtagtcccg	6037
Db	619	ccatttgaggaatgcagaanaagacacccctgcagcttagcagattgccaattcaaccagagaccgc	678
QY	6038	cagaccccaaatctgtccacaatgtgtaactcccaagatctgttaggcaccaagtgaaaccag	6097
Db	679	ctgtgtgcaaaagatgtcttcagatgtgctgtcgaagagcctctgtggagctactgtgaataaccag	738
QY	6098	ggcctttgggaagtgtgccaggtttcttcgtctgtaaatcctagtgaccacaagctcttca	6157
Db	739	gaccacttggaaagtgtgcccaagctgttcttggcttgaattcttcgtatccaataactctatc	798
QY	6158	gacctctataaactctgcactctgcttctaagaatttctctaaagggtgtagaagtgcct	6217
Db	799	gacctctaaacaagtgtgaagttctgctcttaagaattctctctgagatgtagttagaagctg	858
QY	6218	taagaaaaaataaagagcttaattgggcggttcaaaaagagatgtaaaaagagatg99ga	6277
Db	859	taggaaaaaacaagcgtctcatcagcgcagcagaccagagggaaatatacagcagagactcaaaa	918
QY	6278	-aaactatctgcgccttaaagagagccctacagccttagatc--acagaagatccctcagt	6333
Db	919	agaactataacaagctaaagaagacactcagycgaatgtagctcagcggagaaattccaagac	978
QY	6334	tatccaagccagatatt	6349
Db	979	tgtacaagccaatatt	994
RESULT 15			
AAS62625			
ID	AAS62625 standard; cDNA: 1834 BP.		
XX	AAS62625;		
XX	14-FEB-2002 (first entry)		
DE	cDNA sequence #412 encoding novel human secreted protein.		
KW	Human secreted protein: hyperproliferative disorder; autoimmune disorder;		
KW	immune deficiency disorder; blood disorder; inflammatory disorder;		
KW	infectious disorder; gene therapy; antimicrobial; hepatotropic;		
KW	immunosuppressive; antrineumatic; ss.		
OS	Homo sapiens.		

XX MO200177291-A2.
PN 18-OCT-2001.
XX 29-MAR-2001; 2001WO-US10485.
PF 06-APR-2000; 2000US-195604P.
XX (GENE) GENETICS INST INC.
XX Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
PI Gulukota K, Graham JR;
DR WPI: 2002-010900/01.
XX New polynucleotides encoding secreted proteins useful for treating e.g.
PT asthma, HIV and Crohn's disease -
PS
XX Claim 1; Page 292-293; 391pp; English.
XX The present invention relates to the isolation of novel cDNA sequences
CC which encode human secreted proteins. The cDNA sequences have been
CC derived from a variety of human tissues. The invention also provides
CC a method for producing proteins from these polynucleotide sequences.
CC The proteins are useful for identifying compounds that modulate their
CC activity and production, and the cell is also useful for identifying
CC compounds that modulate expression of the polynucleotide sequences
CC encoding the secreted proteins. The sequences of the invention are
CC useful for treating diseases such as hyperproliferative disorders
CC (e.g. cancer), immune deficiency disorders (e.g. severe combined
CC immunodeficiency (SCID)), autoimmune disorders (e.g. multiple
CC sclerosis), blood disorders (e.g. thrombocytopenia), inflammatory
CC disorders (e.g. arthritis) and infectious disorders (e.g. hepatitis).
CC The polynucleotide sequences of the invention are also useful in gene
CC therapy. AAS62214-AAS62838 represent the cDNA sequences of the
CC invention that encode for novel human secreted proteins.
XX
SO Sequence 1834 BP; 502 A; 442 C; 480 G; 410 T; 0 other;

Query Match 5.4%; Score 370; DB 24; Length 1834;
Best Local Similarity 55.0%; Pred. No. 1.1e-82;
Matches 843; Conservative 0; Mismatches 670; Indels 21; Gaps 5;

QY 173 ttacggaacatagatccagtgatgttggaaagtattactctactctcttgctg 232
DB 20 ttatgacctgtggagccagtgacttgaaggaactctgtgacacacccgaaacagcc 79
QY 233 tggatcttgaccttaacggaatttgatgaattccctccagatgatatgagtttt 292
DB 80 tggatgtgagcttgcccgagagctcgagggaacttccagtgagcaacttgagcggtgt 139
QY 293 atagtcctcggaactgcagaactctgtttcagctgtactctgaagaag--tgaatg 349
DB 140 tcacgcaagaagatgtgagactttgcagccctcttcgacggaaggaagggttgactg 199
QY 350 atccaatgttgaagactatagaagaagtatacagaagacgagcgcaattgtcatcagaa 409
DB 200 accctatgctcaggaactgtgtcagacacatccgtgagtggttaacgcygaaccgga 259
QY 410 aatataataatggaagaagatttaaccataatattagataataacaagaagaagc 469
DB 260 aaacacaagaagatccagaatctgtgctttaaaagacttgatctcgaaagaatttc 319
QY 470 aaaaaggttgcacaagaagttttgaaactgtatgaagctccagatgccaagctacc 529
DB 320 acaagacgcttcgaaacagactttgagtcgga-----aaccttgaggtgcagtgaac 373
QY 530 aggatataaagtatgacctaaagaagcgttcaatgtcatatagatatacccaagggtta 589
DB 374 ccgtgtcctcagcagggcccgccacctaacaagtgctgtgagcagtgcttggaagagcc 433

QY 590 gctgggacctgtagtattcttgacttgaanaaatcaactccctgatgtcttgcccaatt 649
DB 434 ccgttaactgctgtgaccttgtagcttcgcagccttgagccttgacaagcgctgagaanaacc 493
QY 650 tacttgatcgaactccaatagaagaatagaccgttagaataatgatgccaaagaaatcaa 709
DB 494 tcttgacgcaagtgagtgccgaggaactttgagaagcagaacagagagcccgagagacca 553
QY 710 accgtcaagaagacttttgcttgatccatccacagatattgagagaggaacacatagaac 769
DB 554 atagcagcgccgagctctcttgcccttaccatccatgagtgacgagagagatgcttgga 613
QY 770 ggtctaggtcttcgtatatacccaagaacatttggtaacagactcttgtaaaatgct 829
DB 614 tagtccagfaccaagaatgtgccaaagaaacacttgccaacagaaatattgtgaagtgc 673
QY 830 tatcaactaaagttagaattgaaattgaaacacatttgcgaattgtgcttatatgatg 889
DB 674 tgaccttgaaagtcagatgaaattgagccctgttgcagacattgccccttaagatg 733
QY 890 tcaaggaagaagaagaatttcagaaacatttatttgaccttaattcttgagcagatga 949
DB 734 ttaagaagaagaaagaatctccagaaatcttcaactgtgacccggaagctcgaccagtcca 793
QY 950 aaggtgttacctgcacatgtaaccaactgtgcacttaactaccctgggaagatlcagca 1009
DB 794 aagattctcgagagctccacagccttcagtgccgcataagtcagcgagatcgcag 853
QY 1010 tttttctcaactatctctcccaagatgttttcttgtaataaagctgaaagaatcc 1069
DB 854 tctctcaagtaacctaccctctccagacactcaactcgtgtagtcaagatgaaagaatcc 913
QY 1070 tacagcaaggaagatcggagagtggtgcagaacatatatgatttccaagaagcgagatg 1129
DB 914 tgcagcggagagagatgtggaactgtgtcagagccctcaacgaagtgatcaagaagtgag 973
QY 1130 ccaaccaaglaaagaanaaactggaagaactgaagatcagaagatcagatgtttccaaa 1189
DB 974 gtggaagaagtgaagaagaagattgaanaaactcaaacactcaagctgatactcttcagc 1033
QY 1190 gacttgggaatatatcgatgccttcttgctgactgcaatccatttaataatgtgtta 1249
DB 1034 gtttgaggaaatlaacggaatgccttctgctgggaccccaataagctatacaagctctt-- 1091
QY 1250 gcaatgctgggaagtttggaaagagattctacagaagtagaatacgaatcgtggaacgaa 1309
DB 1092 -caatgtctccaccccttgagaagggaggttaactgattgga-----ctctgtgttgga 1144
QY 1310 aaggtcttggtcagagagaggaatctagatattgttcgacagatcaacttgaagaaga 1369
DB 1145 gaagctcagtggtggaacgagagacatltggcccaactctagaagagcttct--tgaagag 1201
QY 1370 caacaagtgtgagatgagctgttaacttgaacgagcttgcgaacgactctcaactga 1429
DB 1202 cccctctcttggaaagaaatgtggttggatccaacttcaaacctccacacttgaagctta 1261
QY 1430 caaattttttaagcaggaagagacgccttaagttagtgaagatctctcaaatccctt 1489
DB 1262 gcaactttttaagcaggaagagagatcgcttagcgatggaagactataatgaattttag 1321
QY 1490 ctgatatagaagaagcactctctgttctaagcgagactaagaacchaltacagctcaagctca 1549
DB 1382 gactggagattctacacagctccagagatcaatcaatctgtctcagactccggaatgtgc 1441
QY 1610 aagtgaagcttaccctcagacagtagatgagacattacagagaagaatcttgaattcccg 1669
DB 1442 ccgtgaaaccccttccgtaaaacccggaacgcccgcacaaagagatttggaaatttccaa 1501
QY 1670 caaggaatgttattgttccaaacactacttaacag 1703

Db 1502 cacgagaagtatatgttcctcacactgtgtacag 1535

Search completed: July 16, 2002, 12:24:50
Job time: 13897 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2002, 08:31:43 ; Search time 116.6 Seconds

(without alignments)
14384.096 Million cell updates/sec

Title: US-09-737-246-1

Perfect score: 6828

Sequence: 1 gtcgcgtgcgcgcagcagc.....ctcgtgcgcgcagcgccctc 6828

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	64.2	0.9	7218	US-08-232-463-14	Sequence 14, Appl
2	45.2	0.7	7218	US-08-232-463-14	Sequence 14, Appl
3	44.2	0.6	2317	US-08-749-522-5	Sequence 5, Appl
4	44.2	0.6	3974	US-08-467-504-3	Sequence 3, Appl
5	41.6	0.6	1850	US-08-617-8608-32	Sequence 32, Appl
6	41.6	0.6	4098	US-08-605-106-4	Sequence 4, Appl
7	41.4	0.6	19124	US-08-487-826B-13	Sequence 13, Appl
8	40.6	0.6	6669	US-09-212-971-5	Sequence 5, Appl
9	40.6	0.6	6669	US-08-800-929A-5	Sequence 5, Appl
10	40.6	0.6	6669	US-09-617-053A-5	Sequence 5, Appl
11	39.6	0.6	860	US-08-998-416-287	Sequence 287, App
12	39.4	0.6	837	US-08-998-416-288	Sequence 288, App
13	39.4	0.6	1850	US-08-617-8608-32	Sequence 32, Appl
14	39.4	0.6	1939	US-07-715-751B-2	Sequence 2, Appl
15	39.4	0.6	4098	US-08-605-106-4	Sequence 4, Appl
16	39.2	0.6	1689	US-07-991-867B-41	Sequence 41, Appl
17	39.2	0.6	1689	US-08-544-332-41	Sequence 41, Appl
18	39.2	0.6	8457	US-07-991-867B-1	Sequence 1, Appl
19	39.2	0.6	8457	US-08-544-332-1	Sequence 1, Appl
20	39.2	0.6	2251	US-08-991-677-11	Sequence 11, Appl
21	38.8	0.6	289	US-09-007-005-17	Sequence 17, Appl
22	38.8	0.6	289	US-09-244-796-17	Sequence 17, Appl
23	38.8	0.6	665	US-08-883-795A-36	Sequence 36, Appl
24	38.8	0.6	5852	US-07-867-106-2	Sequence 2, Appl
25	38.6	0.6	658	US-08-998-416-595	Sequence 595, App
26	38.6	0.6	3047	US-07-927-851-2	Sequence 2, Appl
27	38.6	0.6	3047	US-08-453-323-2	Sequence 2, Appl

C 28	38.6	0.6	3047	1	US-08-440-520-2	Sequence 2, Appl
C 29	38.6	0.6	19124	2	US-08-487-826B-13	Sequence 13, Appl
C 30	38.4	0.6	289	4	US-09-007-005-17	Sequence 17, Appl
C 31	38.4	0.6	289	4	US-09-244-796-17	Sequence 17, Appl
C 32	38.2	0.6	662	4	US-08-936-165A-142	Sequence 142, App
C 33	38.2	0.6	2428	2	US-08-849-536A-3	Sequence 3, Appl
C 34	37.6	0.6	522	1	US-08-558-735-7	Sequence 7, Appl
C 35	37.6	0.6	522	4	US-08-906-480-7	Sequence 7, Appl
C 36	37.6	0.6	4612	2	US-08-447-031A-8	Sequence 8, Appl
C 37	37.4	0.5	636	4	US-08-998-416-1137	Sequence 1137, Ap
C 38	37.4	0.5	1107	1	US-08-165-038-3	Sequence 3, Appl
C 39	37.4	0.5	1107	1	US-08-621-081A-3	Sequence 3, Appl
C 40	37.4	0.5	1107	2	US-08-876-781-3	Sequence 3, Appl
C 41	37.4	0.5	1956	4	US-08-559-896B-1	Sequence 1, Appl
C 42	37.4	0.5	2110	4	US-09-419-459-1	Sequence 1, Appl
C 43	37.2	0.5	3618	1	US-07-872-678A-36	Sequence 36, Appl
C 44	37	0.5	2989	1	US-07-927-851-1	Sequence 1, Appl
C 45	37	0.5	2989	1	US-08-453-323-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SRO ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-F1s
US-08-232-463-14

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PT9gpt-F15
US-08-232-463-14

Query Match          0.7%; Score 45.2; DB 1; Length 7218;
Best Local Similarity 6.8%; Pred.No.0.048;
Matches 29; Conservative 211; Mismatches 184; Indels 0; Gaps 0.

Qy 3202 cagaagacagacagaatggttgagagactcaatacaagccttcactcttctcaatgat 3261
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1012 CATACGCTCAGAGATTATTCGAGCTTGCGTGCAGGTCGAGGAGAGCTTCGATATTTT 1071

Qy 3262 cgtgtctcgttatacgacagagagattgttttagccttataaagtcctgtataacag 3321
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1072 YYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYY 1131

Qy 3322 gtgtcttcaagcttactcaatcaccgaatcccaatgttctgtctccttagagctgat 3381
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1132 YYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYY 1191

Qy 3382 ttctcagatcatcctgcagtcagtcagcactatgttacttaacttaccctgcagctta 3441
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1192 YYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYY 1251

Qy 3442 ctactcacactgcatactccacacactctcgtttctctgcacacatcagaagttctgga 3501
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1252 YYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYY 1311

Qy 3502 ttctcagaaatgtcaagaccacaagaatgtcaaatatglttgaattatcogtccttc 3561
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1312 YYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYY 1371

Qy 3562 cgcacacagcatatttgacgagacttgtttaacagagcgtgcattcttagacct 3621
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1372 YYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYY 1431

Qy 3622 gatc 3625
      : |
Db 1432 YYYG 1435

RESULT 3
US-08-749-522-5
; Sequence 5, Application US/08749522
; Patent No. 6096950
; GENERAL INFORMATION:
; APPLICANT: John, Maliyakal
; TITLE OF INVENTION: FTBR-SPECIFIC PROMOTERS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles & Brady
; STREET: 411 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: WI
; COUNTRY: U.S.A.
; ZIP: 53202-4497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/749,522
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Baker, Jean C.

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

? ORIGINAL_SOURCE:
? ORGANISM: Cuphea lanceolata
? IMMEDIATE_SOURCE:
? LIBRARY: genomic Lambda FIX II
? CLONE: CITE91
?
? FEATURE:
? NAME/KEY: CAAT-Signal
? LOCATION: 1428..1432
? FEATURE:
? NAME/KEY: TATA-Signal
? LOCATION: 1553..1556
?
? FEATURE:
? NAME/KEY: Transcription start
? LOCATION: 1585
?
? FEATURE:
? NAME/KEY: Leguminbox
? LOCATION: 1642..1657
?
? FEATURE:
? NAME/KEY: Startcodon
? LOCATION: 1797..1799
?
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1797..1850
?
? US-08-617-860B-32

```

Query Match	0.68;	Score 41.6;	DB 3;	Length 1850;
Best Local Similarity	48.7%;	Pred. No. 0.21;		
Matches 113;	Conservative	0;	Mismatches 119;	Indels 0;
			Gaps	0;

QY	6542	aacatcaaatgtagattgttgaacggcttgaaactcagtcgtatggttccatagtgcgg	6601
Db	450	ATATATATAAATTCTCCGGTATCTCCAACMAAATTTTTCATATTTCCAAATAATTCAA	391
QY	6602	taacaagcgttatctttaagaacatttcaatgaactcaaaggtaacatacatattacca	6661
Db	390	AAAATATTAACACTTTAATATATATTTTTTAAACCAMAAATATTTTAAAAATTCACA	331
QY	6662	ttaattaccatagctaaggttaaataattatccaattgaagtcgfatltttaact	6721
Db	330	ATATATTTAAAAAATTAACCAATCTAATTTTAAAAATATTTTAAAAAACATAATT	271
QY	6722	atatcacacattatagatcatcttiggaccaccttaagttaagtgaatagctt	6773
Db	270	TATATATTTTAAAAATATTTTTTTTAAAAGAATTTAAAAATATATTTTAAATTT	219

RESULT 6
US-08-605-106-4/c

```

; Sequence 4, Application US/08605106
; Patent No. 5910631

```

; GENERAL INFORMATION:

```

; APPLICANT: Topfer, R.
; APPLICANT: Martini, N.

```

APPLICANT: Schell, J.
TITLE OF INVENTION: MEDIUM CHAIN-SPECIFIC THIOESTERS

```

; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:

```

ADDRESS: Schwegman, Lundberg, Woessner & Kluth, P.A.
STREET: P.O. Box 2938

CITY: Minneapolis
STATE: MN

COUNTRY: USA
ZIP: 55402

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

```

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

```

; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:

```

/ CURRENT FILING DATE: 08/08/2008
 ; APPLICATION NUMBER: US/08/605,106
 ; FILING DATE: 03-SEPT-1996

CLASSIFICATION: 800

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/02935
FILING DATE: 01-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 235,001US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-373-6900
TELEFAX: 612-339-3061
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 4098 Base pairs
TYPE: nucleic acid
STRANDEDNESS: double stranded
TOPOLOGY: linear
MOLECULE TYPE: : DNS (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Cuphea lanceolata
IMMEDIATE SOURCE:
LIBRARY: genomic Lambda FIX II
CLONE: ClTE91
FEATURE:
NAME/KEY: CDS
LOCATION: join(1797..2294, 2658..2791, 2898..3011, 3132
LOCATION: ..3303, 3391..3459, 3672..3941)
FEATURE:
NAME/KEY: Startcodon
LOCATION: 1797..1799
FEATURE:
NAME/KEY: exon II
LOCATION: 1787..2294
FEATURE:
NAME/KEY: intron II
LOCATION: 2295..2657
FEATURE:
NAME/KEY: exon III
LOCATION: 2658..2791
FEATURE:
NAME/KEY: intron III
LOCATION: 2792..2897
FEATURE:
NAME/KEY: exon IV
LOCATION: 2898..3011
FEATURE:
NAME/KEY: intron IV
LOCATION: 3012..3131
FEATURE:
NAME/KEY: exon V
LOCATION: 3132..3303
FEATURE:
NAME/KEY: intron V
LOCATION: 3304..3390
FEATURE:
NAME/KEY: exon VI
LOCATION: 3391..3459
FEATURE:
NAME/KEY: intron VI
LOCATION: 3460..3671
FEATURE:
NAME/KEY: exon VII
LOCATION: 3672..3941
FEATURE:
NAME/KEY: Stopcodon
LOCATION: 3942..3944
US-08-605-106-4
Query Match 0.6%; Score 41.6; DB 2; Length 4098
Best Local Similarity 48.7%; Pred. No. 0.34;

```


RESULT 7
US-08-487-826B-13
; Sequence 13, Application US/08487826B

Query Match	0.68;	Score 41.4;	DB 2;	Length 19124;
-------------	-------	-------------	-------	---------------

317 ttgttcagctcttacctgaagaagtgaatcgtatccacattgttagaagctctataaaga 376
Best Local Similarity 52.0%; Pred. No. 1;
Matches 93; Conservative 0; Mismatches 86; Indels 0; Gaps 0.

RESULT 8
HS-09-212-971-5/c

```

Sequence 5 Application US/09212971B
Patent No. 6107041
GENERAL INFORMATION:
APPLICANT: Kornelink, Robert G
APPLICANT: Mackenzie, Alexander E
APPLICANT: Liston, Peter
APPLICANT: Baird, Stephen
APPLICANT: Tsang, Benjamin K
APPLICANT: Pratt, Christine
TITLE OF INVENTION: DETECTION AND MODULATION OF IAPS AND
TITLE OF INVENTION: NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVES
TITLE OF INVENTION: DISEASE
FILE REFERENCE: 07891/009002
CURRENT APPLICATION NUMBER: US/09/212.971B
CURRENT FILING DATE: 1998-12-16
EARLIER APPLICATION NUMBER: 60/017.354
EARLIER FILING DATE: 1996-04-26
EARLIER APPLICATION NUMBER: 60/030.590
EARLIER FILING DATE: 1996-11-14
EARLIER APPLICATION NUMBER: 08/800.929
EARLIER FILING DATE: 1997-02-13
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 5
LENGTH: 6669
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variation
LOCATION: (3677)..(3951)
OTHER INFORMATION: n can be any nucleotide
US-09-212-971-5

```

Query Match 0.68; Score 40.6; DB 3; Length 6669;

Best Local Similarity 49.8%; Pred. No. 0.87;
Matches 103; Conservative 0; Mismatches 104; Indels 0; Gaps 0

Accession	Sequence	Position
Oy	6575 tcaaggcagaggttcctaaagttccggtaacaaagcgttactctttaaagacttcaatg	6633
Oy	2425 TCATAGCTCTTTCACGGGGGACGGAGAACATCACATCTCATTTACCAACATTTATTC	2366
Oy	6635 actcaaggtaacactatatacttaacattattatcacatagctaaggttaaaatttat	6694
Oy	2385 CCTCAAGGTAAAGTAATTTGATCTAATCTTAATTTCTTTCCCGGCCAAAATTTTAC	2306
Oy	6695 tcaacttaagtcgtlactttttaaattatatacaccaattatagatcattcttgagccat	6753
Oy	2305 ACATGTTTAAATAAATCTTCTTTGATATTCACGCTCTTTTAAATTAATTTCTAAGTTT	2248
Oy	6755 tttaaagtgaatgcctattttaa	6781
Oy	2245 CTTCATTTCTTAAATGTAACTTTCA	2219

RESULT 9

US-08-800-929A-5/c
; Sequence 5, Application US/08800929A

; Patent No. 6133437

```

; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G
; APPLICANT: Mackenzie, Alexander E
; APPLICANT: Liston, Peter
; APPLICANT: Baird, Stephen
; APPLICANT: Tsang, Benjamin K
; APPLICANT: Pratt, Christine
; TITLE OF INVENTION: DETECTION AND MODULATION OF
; TITLE OF INVENTION: IAPS AND NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
; TITLE OF INVENTION: DISEASE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Ebling LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,929A
; FILING DATE: 13-FEB-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,590
; FILING DATE: 14-NOV-1996
; APPLICATION NUMBER: 60/017,354
; FILING DATE: 26-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bieker-Brady, Kristina
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 07891/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
;
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-800-929A-5
;
Query Match          0.6%; Score 40.6; DB 3; Length 6669;
Best Local Similarity 49.8%; Pred. No. 0.87;
Matches 103; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
;
QY 6575 tcattgctatggtttcctaatttcggttaacaagctgtatcttctaagaacatttaag 6634
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2425 TCAATGCTTTGTTCCAGGTGGCAGGAGAAACATCATCATCTTACCAACAATATTTC 2366
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 6635 actcaaaagtcacataacattaccattatataccataagctaaagttaaaaattat 6694
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2365 CCTCAAAAGTAAAGTAATGTAATCTCAATATTCCTTCCCTGGCCAAAATTTTAC 2306
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 6695 tcacttaagtcgtatttttaattatataccattatataccattatagattttggaccat 6754
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2305 ACATTGTATATAAATCTCTTTGTATTCATCATCTGCTTATATTAATTTCTAGGTT 2246
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 6755 tttaaatgtaagtaagcttattttaa 6781
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2245 CTTCATTTCCTAAATGTAACTTCAA 2219
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
;
RESULT 10
US-09-617-053A-5/c
```

```

; Sequence 5, Application US/09617053A
; Patent No. 6300492
;
; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G
; APPLICANT: Mackenzie, Alexander E
; APPLICANT: Liston, Peter
; APPLICANT: Baird, Stephen
; APPLICANT: Tsang, Benjamin K
; APPLICANT: Pratt, Christine
; TITLE OF INVENTION: DETECTION AND MODULATION OF IAPS AND
; TITLE OF INVENTION: NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
; TITLE OF INVENTION: DISEASE
; FILE REFERENCE: 07891/009003
; CURRENT APPLICATION NUMBER: US/09/617,053A
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/800,929
; PRIOR FILING DATE: 1997-02-13
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 6669
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (3677)...(3951)
; OTHER INFORMATION: n can be any nucleotide
;
US-09-617-053A-5
```

```

Query Match          0.6%; Score 40.6; DB 4; Length 6669;
Best Local Similarity 49.8%; Pred. No. 0.87;
Matches 103; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
;
QY 6575 tcattgctatggtttcctaatttcggttaacaagctgtatcttctaagaacatttaag 6634
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2425 TCAATGCTTTGTTCCAGGTGGCAGGAGAAACATCATCATCTTACCAACAATATTTC 2366
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 6635 actcaaaagtcacataacattaccattatataccataagctaaagttaaaaattat 6694
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2365 CCTCAAAAGTAAAGTAATGTAATCTCAATATTCCTTCCCTGGCCAAAATTTTAC 2306
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 6695 tcacttaagtcgtatttttaattatataccattatataccattatagattttggaccat 6754
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2305 ACATTGTATATAAATCTCTTTGTATTCATCATCTGCTTATATTAATTTCTAGGTT 2246
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 6755 tttaaatgtaagtaagcttattttaa 6781
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2245 CTTCATTTCCTAAATGTAACTTCAA 2219
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
;
RESULT 11
US-08-998-416-287
; Sequence 287, Application US/08998416
; Patent No. 6239264
;
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reibischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
```

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 287:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 860 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1240UP
;
; US-08-998-416-287
```

```
Query Match          0.6%; Score 39.6; DB 4; Length 860;
Best Local Similarity 56.0%; Pred. No. 0.46;
Matches 75; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
```

```
Oy 6662 ttattatcacatgactaaggttaaaattattcacttgaagtcgatttttaattc 6721
    ||||| ||| ||| ||||| ||| ||| ||||| ||| ||| |||||
Db 443 TTATTAATAAATAAGCTATATTATTAATTAATAATTAATAATTAATAATAT 502
Oy 6722 atatcacattatagatttccttgaccatttaattagtaagtcctatttaa 6781
    ||||| ||| ||| ||||| ||| ||| ||||| ||| ||| |||||
Db 503 CTATAGTAATTTCTTATTATTATTATTAACATTTTAAATGTTTAAATAGTA 562
Oy 6782 gtactacaaaata 6795
    ||| |||||
Db 563 AFAACAATTAAATA 576
```

```
RESULT 12
US-08-998-416-288
; Sequence 288, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Redischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSTYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 288:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 837 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1241RP
;
; US-08-998-416-288
```

```
Query Match          0.6%; Score 39.4; DB 4; Length 837;
Best Local Similarity 43.9%; Pred. No. 0.52;
Matches 169; Conservative 0; Mismatches 216; Indels 0; Gaps 0;
```

```
Oy 6396 aaatgacatcctcaactgaatgacattgttatcactcgaagaagcattatca 6455
    ||||| ||| ||| ||||| ||| ||| ||||| ||| ||| |||||
Db 187 AAAAATAATATTTAATGATGATCTATTAGCTATGTTCAATTTAAATAGTATTA 246
Oy 6456 acatcgagtgtaaaagatctatggaaacacacatggaatgcgaattatata 6515
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 247 AAATATTATTAGATATTATTATTCTTTAAATAATTAATTAATGATATCAATAATTA 306
Oy 6516 ttcatgagaatgcagctggccaaagaataatcagaatgtaagttgtaacgcttgaat 6575
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 307 ATATATTATTATTAAATGTTATTTAATAAATAATTAATTAATTAATTAATTAAT 366
Oy 6576 catgacctggttctcaatgctcgggtaacaagctgtatccttlaagaactttaa 6635
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 367 TATTTAATATTGTAATATTATTATTATTATTAATATCTATTTTAATAATATATGTT 426
Oy 6636 ctcaaggtacactatacatattacattatcacatagctaaaggttaaaattatc 6695
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 427 GATTATTATTATTAAATCTTTTAAGAATATTATTAAATAATTAATTTAACTTAATT 486
Oy 6696 cacttgaagtcgatttttlaattatataccacttataagattcatttgaccact 6755
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 487 TCTTATTATTAAATTTTAAATTAATTAATAAATAATTAATTAATTAATTAATTAAT 546
Oy 6756 ttaaatgtaatgctattttaa 6780
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 547 TTAATTAATTAATTAATTAATTAATTA 571
```

```
RESULT 13
US-08-617-860B-32
; Sequence 32, Application US/08617860B
; Patent No. 6133506
; GENERAL INFORMATION:
; APPLICANT: Tytfer, R., Bautor, J., Bothmann, H., Filasak, E.,
; APPLICANT: Hvirike-Grandpierre, C., Klein, B., Martini, N.,
; APPLICANT: Miller, A., Schulte, W., Voeltz, M., Walek, J.,
; APPLICANT: Schell, J.
; TITLE OF INVENTION: Promoters
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
```



```

RESULT 15
US-08-605-106-4
; Sequence 4, Application US/08605106
; Patent No. 5910631
;
; GENERAL INFORMATION:
; APPLICANT: Topfer, R.
; APPLICANT: Martini, N.
; APPLICANT: Schell, J.
;
; TITLE OF INVENTION: MEDIUM CHAIN-SPECIFIC THIOESTERS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,106
; FILING DATE: 23-SEPT-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/02935
; FILING DATE: 01-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 235,001US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
;
; TELEX:
;
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4098 Base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double stranded
; TOPOLOGY: linear
; MOLECULE TYPE: : DNS (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Cuphea lanceolata
; IMMEDIATE SOURCE:
; LIBRARY: genomic Lambda FIX II
; CLONE: C1TEg1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1797..2294, 2658..2791, 2898..3011, 31322
; LOCATION: ..3305, 3391..3459, 3672..3941)
; FEATURE:
; NAME/KEY: Startcodon
; LOCATION: 1797..1799
; FEATURE:
; NAME/KEY: exon II
; LOCATION: 1787..2294
; FEATURE:
; NAME/KEY: intron II
; LOCATION: 2295..2657
; FEATURE:
; NAME/KEY: exon III
; LOCATION: 2658..2791
; FEATURE:
; NAME/KEY: intron III
; LOCATION: 2792..2897
; FEATURE:
; NAME/KEY: exon IV
; LOCATION: 2898..3011

```

```

? FEATURE: intron IV
? NAME/KEY: intron IV
? LOCATION: 3012..3131
? FEATURE:
? NAME/KEY: exon V
? LOCATION: 3132..3303
? FEATURE:
? NAME/KEY: intron V
? LOCATION: 3304..3390
? FEATURE:
? NAME/KEY: exon VI
? LOCATION: 3391..3459
? FEATURE:
? NAME/KEY: intron VI
? LOCATION: 3460..3671
? FEATURE:
? NAME/KEY: exon VII
? LOCATION: 3672..3941
? FEATURE:
? NAME/KEY: Stopcodon
? LOCATION: 3942..3944
?
? US-08-605-106-4

```

	Query Match	0.6%	Score 39.4;	DB 2;	Length 4098;	
	Best Local Similarity	50.3%;	Pred. No. 1.4;			
	Matches	97;	Conservative	0;	Mismatches	96; Indels 0; Gaps 0;
QY	6614	aacctttagacattttaagcaactcaaaggtaacatacacatctaccattattaca	6673			
		" " " " " " " " " "				
Db	209	AATATTTTTTTAAATAATATTTTAATTCCTTTAAAAAATAATTTTTTAAATTA	268			
QY	6674	tagctaagtgttaaataattatccacttaagtcgatatcttttaattataatacactt	6733			
		" " " " " " " " " "				
Db	269	TAAAATTAGTGTTTAAATAATTTTTTAAARATTTTAAATTAACGTTTTTTTGTATTTTAAAT	328			
QY	6734	alagaatcatcttggaaccactttaaatgltgatgaatgcattttaaaagtlactaaaaa	6793			
		" " " " " " " " " "				
Db	329	ATGGTCGAATTTTTTAAATAATTTTTTGGTTTAAANAATATFTTAAAGAATTTTAAATAT	388			
QY	6794	tatgtgaatgctt	6806			
		" " " " " "				
Db	389	TTTTTGGAATTTTT	401			

Search completed: July 16, 2002, 12:12:08
Job time: 13225 sec

```

/codon_start=1
/protein_id="CAC43726.1"
/db_xref="GI:14597912"
/translation="MAERARAFQKISRVAAEVKQISGQYSGSPOLIKININIGNTS
HTTAPLEAVDPVDLEDYLIITHPLAVDSGLRLIEPPDIEEVVSPRCRTLVSA
VVESEMPHVCDCIRSTEDMAIVIRKYNHLDGTGDKREKQNTGDKPKQVFE
DEARDNSYQDDODDLKRRMSIDTPEERWAGSIFDLKNDIPALLPNTLIDRPN
IDRONDDQKSNRHKELFALHSPDEEPIRSLPDKPKHFORILVLCISKEPI
ETEPFASIALYDVEKKKISENFYFDLSFROMGKLRPHYPAATITTLARSAFST
YPSODVEFLVLEKYLOOGDICEAEPMTREKDATKNKELKLSQADFCORLG
KYRMPFATAIHLNMTVSAGSLEDSTEVEITSGERKGSBERNSIVGRSLEET
TSGDACNLTSFPAATLVTFNFKQEGRLDEDLVFLDMRPSSVLRRLRTIQ
LKIDISAPENPHCLPPELLQVLYLPSRVRPRLIEFARVYVYPPNTYRLIYI
YPOSLEFNKROGSARNLIVKQFMGDEPSNAMYIFGKSSCESEKATVAAVYHR
SPDEHEIKVPLATIDHHLLEFHYVSCQOONPLETPGVGTWYPMONRLKT
GOFCLPVLEKPPQAYSYLSPVELPCKMGKDNKGVNENYVAVSSITHODPLDKF
FALVNAIDENHFPFRIGDMRMENNELKSSISALNSOLEVAVFHLDLKLL
LVIRPVIAGQIVNLGQSFEMASIIINRLKRLNLEGNHDOGRNSLASIHYVRLP
NTYPPSSSPGGGLGSGVHYATMARSAVPRASLLNRSLSNSNPDISGTPSPDE
VRSIIGSKGLDRSNSWVTGPKAAPGWSNPSPAESTOAMDRSCNMSHETESFL
OTLGRLPYTKLFHEBELALQWVCGSVRESALQAMFFELAMKSVHHLXENDKE
APRSRPEREMDIIALVSTIASDIYSRQKDIEMERLNTSLAFNLNLSYMDNG
FVPSLIKCYKQVSKSLSLPNSVLSLRDLRLTICSHHYVTLNLPCLSLTPPS
PSPVSATSSQSGFSTINVDQKIANNEELSVPRQOXYLAGLVLELAVILDPDAG
LEGLHKRYINWNLHSHSDPRYSDPOLKARVAMLYLIGIMETVQLDFTET
HNORGRPICITDYESGSGMSIQTAYAMLAGSVOLRPGSFLTSTSGRHTTE
SAESSRLICILWLWKNADETVLQKMTDLSVQLRLIDLILYICVSEHYGKRYF
ERMSLTPEKSKDMRAKLEPAIILISIGARQEMVRSQGLERSPSGSARFQENLNR
KDNTHMKNTEKLDKRAELIHEALILIGNLATEANLILITLLETIVQVSTESKEI
LGVLVLYLHSMACNOSAVYLOHCFATQRLVLSFPLLEBEETVQADCLRLRHC
SSSIGTIRSHPSALYILMRQNEPILGNMFARVKVQWMSLSHEDQSONNEFLRS
LKITLITAEEDLELRETFPPQVDLFINHMLISDVKNKEHODEDEMILDMYRA
KGYOTSPRLTLONMAGKHSERSNHAEACQVLSAALVAEYTMLEDKTYLPGVC
TFONISSVLEESAVSDVYSPDEEGTCSKQFTEGSLVLIQLQAAASFPMAGHYEV
NEVYKVLPIHEANRAKUSTITGKIQDAPSKYVHOSTMERFRTGTFVYGTKE
GDLPDEEVYKEPATITRLAETSHRLESGFERGEDVEVEIKQSNPDVKCKLDPNAY
IQITVBERYEDYEMKDRITFYEDKNYLRFRMYCTPTLDRGNAHGBELHEDFKERTIT
TSHAPEYIKRVTNTHKEEILIPILIEVALIDEMOKTLEAFTHODPADKMLQMVLO
GSGVTYVNOGLAEVAQVLEIPSDPRLFRHNHKLRLCFDPLRCDADARKKSLIG
PVQKEYOREIKLSSP"
BASE COUNT      2082 a      1384 c      1432 g      1930 t
ORIGIN
Query Match      100.0%; Score 6828; DB 6; Length 6828;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 361 AGAAGCTGTATTAAGAGATTATACGAAGACGTGGCAATTTGTCAAGAAAATATCATATAA
421 ttgggaacaggaatttaa tcccaataca ttagataaacaagaagaagccaagaagtttg
421 ttgggaacaggaatttaa tcccaataca ttagataaacaagaagaagccaagaagtttg
481 ccaaaacaagtttttgaa tcttgatgaag ctcagaatgcaacagctlaacagatgatcaa
481 ccaaaacaagtttttgaa tcttgatgaag ctcagaatgcaacagctlaacagatgatcaa
481 ccaaaacaagtttttgaa tcttgatgaag ctcagaatgcaacagctlaacagatgatcaa
541 gataactiaaagaagctca atgtaataa tagataaaccacaagggttagcttgagctgt
541 gataactiaaagaagctca atgtaataa tagataaaccacaagggttagcttgagctgt
601 agtactttgacttgaaaaa ttaacattca cttcagatgcttgcttcccaattacttgatcga
601 agtactttgacttgaaaaa ttaacattca cttcagatgcttgcttcccaattacttgatcga
661 actccaaatgaagaagaat agaacccg tgcgaatgatagcaaaagaaatccaacgtaacaa
661 actccaaatgaagaagaat agaacccg tgcgaatgatagcaaaagaaatccaacgtaacaa
721 gaaacttttgcttgcaatc ccatcaca cagatgagaagaacaaatagaacgcttagttt
721 gaaacttttgcttgcaatc ccatcaca cagatgagaagaacaaatagaacgcttagttt
781 cctgataatacccaagaaca ttttgctcaa agactcttgtaaatgctlaatacaccaga
781 cctgataatacccaagaaca ttttgctcaa agactcttgtaaatgctlaatacaccaga
841 ttggaattgaattgaaccca ttttgcaag ttttgcttataatgataatgcaagaagaag
841 ttggaattgaattgaaccca ttttgcaag ttttgcttataatgataatgcaagaagaag
901 aaaaagattccaagaacact ttaatttg aacttaattcgaagcagatgaaaggttgata
901 aaaaagattccaagaacact ttaatttg aacttaattcgaagcagatgaaaggttgata
961 gctccacatgataccaccgt cgcacata ctaaccctggaagaatcagaacattttctatc
961 gctccacatgataccaccgt cgcacata ctaaccctggaagaatcagaacattttctatc
1021 acttaaccctccaagaatg ttttcttg taataaagctagaagaagctcttaagaagaaga
1021 acttaaccctccaagaatg ttttcttg taataaagctagaagaagctctcttaagaagaaga
1081 gacattggaagatgtgca gaacacata tatgatttccaagaagcagatgccacaagaat
1081 gacattggaagatgtgca gaacacata tatgatttccaagaagcagatgccacaagaat
1141 aaagaagaactgagaagaac tgaagag tcaagcagatcagtttgcacaagacttgggaaa
1141 aaagaagaactgagaagaac tgaagag tcaagcagatcagtttgcacaagacttgggaaa
1201 tatcgatgacctttgcttg agactgac tcaatcatttaataaattgttgacagctctggg
1201 tatcgatgacctttgcttg agactgac tcaatcatttaataaattgttgacagctctggg
1261 agtttgaaagagatctca caagaagtag aatcagtaactcagtgagaagcgaagaagctctgg
1261 agtttgaaagagatctca caagaagtag aatcagtaactcagtgagaagcgaagaagctctgg
1321 agtttgaaagagatctca caagaagtag aatcagtaactcagtgagaagcgaagaagctctgg
1321 agtttgaaagagatctca caagaagtag aatcagtaactcagtgagaagcgaagaagctctgg
1381 gatgagctctgtaacttg acagagcttg accagcttaactcagtgagaagcgaagaagctctgg
1381 gatgagctctgtaacttg acagagcttg accagcttaactcagtgagaagcgaagaagctctgg
1441 gatgagctctgtaacttg acagagcttg accagcttaactcagtgagaagcgaagaagctctgg
1441 gatgagctctgtaacttg acagagcttg accagcttaactcagtgagaagcgaagaagctctgg
1500 aagcagaagaagcagccgt taagtgaaga tctctcaaaatctctgcgtgataatgaga

```



```
Db 1441 AAGCAGAGAGAGACCGCTTAAAGTATGATGAAGATCTCAACAATTCTTGGTGAATGAGA 1500
Oy 1501 aggcacatctctgtccttaaggcgactaagaacctatacagctcgaagctcaagaatagaatt 1560
Db 1501 AGGCATCTTCTGCTTACGCGGACTAAGACCTATTACACTCAGCGCAAGATAGACATT 1560
Oy 1561 tctccgacacttgaaatcccccatatctgcctaactccgagagctgcgaagtgaagctc 1620
Db 1561 TCTCCGACCTTGAAATCCCATATTATGCTTAACCTCCGAGCTGCTTCAAGTGAACCTT 1620
Oy 1621 tacctgacagtagagtagaagcttaacagagaaatccttagagttcccgcaaggagatgtt 1680
Db 1621 TACCCTGACAGTAGAGTTAGACTACACAGAAATCTTAGAGTTCCCGCAAGGAGATGT 1680
Oy 1681 tatgttccaacactacttaagaagaatctctctacatatacctctgaagttcttaattt 1740
Db 1681 TATGTCCAAACACTACTACTACAGAAATCTTCTACATATACCTCCAGAGCTTAAATTTT 1740
Oy 1741 gccaatcgtaaggtctcgtcagaaataaataacagtgaaagccagttatgtatgagag 1800
Db 1741 GCCAATCGTCAAGGTCTGCTGAGAAATATACAGTGAAGCCAGTTTATGTATGAGAG 1800
Oy 1801 gatccaagaatgcacatgcggttaactcttggttaaatcagctgttcaagaatttcaag 1860
Db 1801 GATCCAAGCAATGCCATGCCGTATCTTTGTAAATCTACCTGTTGAGAAATTTTCAAG 1860
Oy 1861 gaagccttaagccgttatcatcaataacaggtctcctgatttctcaatgaagaataag 1920
Db 1861 GAAGCCTTACAGCCGTATGATATCATACAGGTCTCCTGATTTTCAAGAAATACAG 1920
Oy 1921 gftaagcttcctgtactacttaactacatcacatcactgcttcttaacttatacatgt 1980
Db 1921 GTTAAAGCTTCCGCTACTTAACTACATCATCATCACTTGCTTTTACTTTTATCATGTT 1980
Oy 1981 agtgtcaacaaacaaataactcctcttgaacacacagttgatatatacatgatacca 2040
Db 1981 AGTGTCAACAAACAAATAACTCTCTTGAACACACAGTTGGATATACATGATGATACA 2040
Oy 2041 atgcttcgaatgagacggttgaagactgacagtttgcttgccaagtctcatgttgaaaa 2100
Db 2041 ATGCTTCAAGAAATGAGCGTGTGAAGACTGACCAAGTTTGCTTGCCAGTCTCATGTGAAA 2100
Oy 2101 ccacacagagctatctcgtactcgtcctcgtaggtctcctcaactcgtcagatgaatggt 2160
Db 2101 CCACACAGAGCTTATCTGTACTGTCTCTGAGGTCTCTACCGGAGTGAATGAGGTA 2160
Oy 2161 gataatcaacaaagtglttttaagtgtgaagttgctgtctgtctcatccatacaca 2220
Db 2161 GATAATCAACAAGTGTTTAAATGTTGAAGTTGTGCTGTTCCGTATATCCATACACAA 2220
Oy 2221 gatccttatcttgacaataatttctgctgtgcaatgctctgagatgaacactgttcca 2280
Db 2221 GATCCTTATCTTGACAAATTTTCTGCTGCTCAATGCTGTGAGTGAACACTGTTCCTCA 2280
Oy 2281 gtcgcaattgaggacatgcgaatcactggaataaacttagaataatgaaatgaaagcagt 2340
Db 2281 GTCCGAAATGGGGACATGCGAATCATGGAATAATTAAGTAAGAAATGAAAGAGCAGT 2340
Oy 2341 attcaagcactgaatcatccagctggaacagatggtccagttctctcatcctctgcta 2400
Db 2341 ATTTCAGCACTGAATTCATCCAGCTGGAACAGTGGTCCGATTTCTTCACTCTTGCTGA 2400
Oy 2401 gataaactgaacttttagttatagacctcctgcatatgctgagcgaataagttaacta 2460
Db 2401 GATAAATCTGATCTTATTAATTAATAGCTCTGTGATATGCTGGCCAAATTAATTAACCTA 2460
Oy 2461 ggtcaagcactctttagaagccatgycatcaatataaactcgaactcacaaaacttggaa 2520
Db 2461 GGTCAAGCACTCTTTTGAACCATGGCATCAATTAATAATCGACTTCACAAAAAATCTGGAA 2520
Oy 2521 ggaatcatgacacagcatgagcgaagaacagcctctgtgcatcatatcatcatatgttttc 2580
Db 2521 GGAAATCATATGACCATGGATGGCAAAAACAGCTTCTTGATCATATATATCATATATGTTTC 2580

Oy 2581 cgcttaacaatacttaacctaatcatcatcaccaggtctcgtgggtgttggagagatca 2640
Db 2581 CGCTTACCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2640
Oy 2641 gtgcattatgccaacaaatgtagatctgcgtgtagacactgcaagccttaatttaactgt 2700
Db 2641 GTGCATTATGCCCACAAATGAGCTAGATCTGAGTGGAGACCTGCAAGCCTTAATTAATGCT 2700
Oy 2701 tctcgaagccttagtaataatgcgaatccagatatatctggagaccccaagttcccaatgat 2760
Db 2701 TCTCGAAGCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2760
Oy 2761 gaagtcgaatcaatcactgagtagaaggttttagatcgtcccaattcctggtttaaact 2820
Db 2761 TATCGAAGCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2820
Oy 2821 ggtgtccaaaagcgtcccatgaggatcccaaccagttccaaagtgcagaatcaacaag 2880
Db 2821 GGTGTCCAAAAGCCTTACCATGAGGATCCCAACCCAGTCCAAAGTGCAGAAATCAACACAG 2880
Oy 2881 gctatgagtcgaagttgtaatcgtatgcttgcacacagagagcgtcaagttcttaaa 2940
Db 2881 GCTATGAGTCAAGTGTATGATGATGCTTGCCACACAGACAGCTCAAGTTCATTACA 2940
Oy 2941 acattacgagacgcttaacaaactaaagccttctcaagagagctggtcttgcagtg 3000
Db 2941 ACATTAACGGGAGAGCTTACCAACTAAAAAGCTTTTACAGAGAGAGCTGCTTGCAAGTGG 3000
Oy 3001 gttgttgcagtgagcagcgttcgggaatcaagctctgcaacaagcgtgtcttctttttaa 3060
Db 3001 GTTGTGTGAGTGGAGCGCTTCCGGAAATCAGCTTTCACCAACACCCGCTGTTTGTGAA 3060
Oy 3061 ttaatggaagaagccttggtgacacattatatttaagataaactttagagcttccaag 3120
Db 3061 TTAATGGAAGAAGCCTTGAGTGGACACCAATTAATTAATTAATTAATTAATTAATTAATTA 3120
Oy 3121 aaaaagcgttcccaagaagcttccatgagatgacattgacagcgtcttgcagacagattgt 3180
Db 3121 AAAAGCTGTTTCCAGAGAGCTTTCATGAGATGACATTGCAAGCTTGTGCAAGCATTTGCT 3180
Oy 3181 agtataatagttcaacagatccagaagacagaaatggttgaagagactcaatacaag 3240
Db 3181 AGTATATATGTTTACAGATTTTCAAGAGGACACAGAAATGTTGAGAGACTCAATTAACAC 3240
Oy 3241 ctgcatcttcttcctcaatgactggtgtgtgtgtatgagacagagattgttcttaagcct 3300
Db 3241 CTGCAATCTTCTCAATGATCTGTTGTGTGTATGAGACAGAGGATTTGTTTATGCCCTT 3300
Oy 3301 ataaagctcgtcctaaacaggtgtcttcaaaagcttaactcaatcaatcccaagtglt 3360
Db 3301 ATAAAGCTCTGCTATAAAGAGGTGTCTTCAAAAGCTTATACATTAACGAATCCAGAGT 3360
Oy 3361 ctggtctccttgaagctgatttctcagaatcatcgtcaatcagatcagactatgttaca 3420
Db 3361 CTGCTGCTCTTGAAGCTGATTTTCTGCAATCATCTGCAATGAGCTATGAGCTATGTTACA 3420
Oy 3421 ttaacttaacctgacacttaacttaactcaactgcaactcaactcaactctgttctct 3480
Db 3421 TTAACCTTACCTTCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3480
Oy 3481 gcaacatcaccagatctctggaatttctcagaatgtaaaagacaaagaatgtgcaaatatg 3540
Db 3481 GCAACATCTCGAGATCTCGGATTTTACAGAAATGTAACAAACAAAGATGCAAAATATG 3540
Oy 3541 ttgaaatcatcgtgcttcccgcaacagcatatttgcagagagactgtgtttaaagag 3600
Db 3541 TTTGAATTAATCCGTCCTTCCGCAACAGCATTAATTTGGACAGACTTGTTTAAACAGAG 3600
Oy 3601 ctgagctcatatttaagacccctgagctgagagagactgtttagatgcaataagaagcatc 3660
Db 3601 CTGAGCTCATTTTAAGACCTGATGCTGAAGAGACTGTTTGGATTGCTAAGAAAGTCAATC 3660
```

QY 3661 aataatgatacaaatctactctccagtcagacgacccgagctactctgacccctag 3720
| | | | |
Db 3661 AATATGATGACACAAATTACTCTCCAGTCACAGCTACAGCCCGGCTACTCTGACCCCTAG 3720
QY 3721 ataaagctcagatgagccatgcttctatctactctgattgattatcaatgaatctga 3780
| | | | |
Db 3721 ATAAAGGCTCGAGTGGCCATGTTGATCTACTCTGATTGATTATCATGGAACCTGTA 3780
QY 3781 cctcagctgatatgtattacagaagaactcaacaatcaagaagaagaaccatttgtatagcc 3840
| | | | |
Db 3781 CCTACGCTGTATGATTTTACGAACCTACATCTACAGAGAGAACCAATTTGATATGCC 3840
QY 3841 actgatgatataagaatgagagcggaaatgataagaacgagacggtttccatgcaatc 3900
| | | | |
Db 3841 ACTGATGATTTATGAAGAGTGAAGCGGAGGTATGATTAAGCCAGACGGTTGCCATGGCAATC 3900
QY 3901 gaaggagacatggttccctcaactaacaaggcctggcagtttccctcctcaagtcacagat 3960
| | | | |
Db 3901 GAGGGACATCGGTCCCTCACTACCAAGGCGCTGGCACTTTCCTCCTCAACGAGT 3960
QY 3961 ggcaggaacaacactactctttcagcagaatcaagtcgaagccttttgtatctgtactt 4020
| | | | |
Db 3961 GGCAGGCAACACACTACTCTTTACAGCAATCAAGTCAAGCCCTTTGATCTGTACTT 4020
QY 4021 tgggttccaaaatgacatgataaagcttctacagaagtggtttacagatctctcagtc 4080
| | | | |
Db 4021 TGGGTCTCAAAAATGACGATGAACAGTCTACAGAAAGTGTTTACAGATCTCTCAGTC 4080
QY 4081 ttgcagctaaacggcctcttagatctgtcttatactctgtgtgtcttcttgatataa 4140
| | | | |
Db 4081 TTGCACTAAACCGGCTATTGATCTGTTATCTGTGTGTTCTTGTGAATATAAA 4140
QY 4141 gggaaaaaagtgtttgaaacgaatgataagcttgacccctttaaagaataaaagacatgata 4200
| | | | |
Db 4141 GGGAAAAAGTGTGTAACGATGATAGTGTGACCTTTAAGAAATCAAAAAGCATATAGA 4200
QY 4201 gcaaaagcttgaagaagaactatctctgggagcaataagtgccagcaagaatgatacggcga 4260
| | | | |
Db 4201 GCAAAAGCTTGAGAGAGCTATTCTTGGGAGCATAGTGCAGGCAAAAGTGTCGCGCGA 4260
QY 4261 agccgagagacagctcgaagaagaagcccatctggaagtgctttggaagtcaagaagaatttg 4320
| | | | |
Db 4261 AGCCGAGAGACAGCTCGAGAGAGAGCCCATCTGGAAGTGCCTTTGGAATCTCAAGAAATTTG 4320
QY 4321 aggtggagaaagaataatgactactggtcgtcaaaaacacagaagaagcttgacaatcaaga 4380
| | | | |
Db 4321 AGGTGGAGGAAGATATGACTACTGGCGTCAAAACACAGAGAAAGCTTGACAAATCAGA 4380
QY 4381 gcaagaattgagacagcgaagcaactgtatgtatggaacactggtctacagaagaacaaactaatc 4440
| | | | |
Db 4381 GCAGAAGATTGACACGAGACACTGATTGATGGAACCTGGCTACAGAGAACCAACTATATC 4440
QY 4441 attttaagatacatagagatgtgtgttccagaccgttctgtaacggaatccaaagaagc 4500
| | | | |
Db 4441 ATTTAGATACATTAAGAGATTGTTGTTCAACACCTTTCTGTAAAGGAATCCAAAGAGAGC 4500
QY 4501 attcttggtgagtgatgataaagtgtactatacaagcattggcctgttaaccaagtctgaact 4560
| | | | |
Db 4501 ATTCTTGAGTGGAGTGTAAAGTGCTACTACACAGCATGSGCTGTAAACCAAGTGCAAGT 4560
QY 4561 tatctacaacacgtttttgttacaagaagaagcctgtgtttccaagtcttctgaactcta 4620
| | | | |
Db 4561 TATTTACAAACCTGTTTGTCTACACAGAGCCCTGGTTTCAAAAGTTCCTGGAACCTTAA 4620
QY 4621 ttctgaagaagagacagacagtgatgtgctgattatgacctcagagcttctcgaactttagc 4680
| | | | |
Db 4621 TTTCGAAGAAGAGACAGAGCTGTGTGATTTATGCTCAGCTTCTCCAGACCTGATG 4680
QY 4681 agtagcatcggtatacaataggttcaacaccagatgctcccttactactaataatgagca 4740
| | | | |
Db 4681 AGTAGCATCGGTACAAATACGCTCACACCCAGTGCTCCCTTACTACTAATGAGGCAA 4740
QY 4741 aactttagagttggaataacttgcagaggttaaaaatgcaagttaccaatgttactatacc 4800
| | | | |

Db 4741 AACTTTAGATTGGGAATTAATTGCGAGGCTPAAATGAGGTACTCAATGTCATATTC 4800
| | | | |
QY 4801 tcccttggtggacatctccagaattttaaagaagaatttctaagacttctctaagact 4860
| | | | |
Db 4801 TCCTTGGTGGCAATCTCAGAAATTTTAATGAAGAAATTTCTTAAGAGCTTCTCAAGACT 4860
QY 4861 atattgacataatgctgaagaagaatctggaatgtagggaaacaacatlltccgtacagtc 4920
| | | | |
Db 4861 ATATTGACATATCTGTAAGAGATCTGAAATGAGGGAACAACATTTCTGTACAGTTC 4920
QY 4921 caagatctggttttcaatctccatgatatgtcttcttgatactgtgaaatgaaagaac 4980
| | | | |
Db 4921 CAGAGATGGTTTTCATATCTCCATATGATTTCTTCTGATACGTGAAATGAAAGAAC 4980
QY 4981 caagagatctctgaagaatgtagtatagttacaagaattgccaaggtttaccagacc 5040
| | | | |
Db 4981 CAGAGATCTCTGAATGTTGATTGATTAATGATACAGAAATGGCAAGGTTTACAGACC 5040
QY 5041 tctccagagcagatgagacctggttgcagaacatggcaggaagcaactcagaacgaagaat 5100
| | | | |
Db 5041 TCTCCAGAGGATTTGACCTGGTTGAGAACTGGCAGAACACACTACAGAACGAACAT 5100
QY 5101 catgctgaagctgcacagtgatctagtcacactcagcagacactgtgtgtgataatttggc 5160
| | | | |
Db 5101 CATGCTGAAGCTGCACAGTGTCTAGTCCACTCAGCAGCATTTGTTGTAATATTGAGC 5160
QY 5161 atgctgaggaacccggaataatcttccgtgagatgataacttcaagaatatctcatct 5220
| | | | |
Db 5161 ATGCTGAGAGCCGGAATAATATCTCTGTGGAGTGTGAACATTTACAGAAATTTCACT 5220
QY 5221 aatgttttgaagaatctcgtgtcagatgatatgtgtatctcccgatgaaagaagtatc 5280
| | | | |
Db 5221 AATGTTTGAAGAAATCTCGGCTCAGATGATGTGTCCAGATGAAGAAAGCTATC 5280
QY 5281 tgcctggaataatcttactctagtcagactgtggaattactctgaaacaagactgct 5340
| | | | |
Db 5281 TGCTTGGAATAATCTTACTGAGTCAAGACTTGGGATTTACTGGAAACACACTGCT 5340
QY 5341 tccctctcctatgctgagcatgtatgaagcagttaatgaaagttacaagtlactatcct 5400
| | | | |
Db 5341 TCCCTTCTATGGCTGGCATGTATGTAAGCAGTAAATGAAGTAACTTATTTCT 5400
QY 5401 attcaagaagcctaactcggatgataagaagaactcacaaatcatgataacttaaga 5460
| | | | |
Db 5401 ATTCAAGAAGCTAATCGGATGCAAAAGAACTATCTCACATTTCAATGTAATTTAAAGAA 5460
QY 5461 gcaatcaagaataatgttcatcagaagtlactggtcgtggaacggatgttggaacctatct 5520
| | | | |
Db 5461 GCATTCAGCAAAATTTGTCATCAGATCTGCTGGAGCGGATGTTGGCACCTATTT 5520
QY 5521 cgttggtgttttataggaaacaaagttcggggaattggatgaacaagaatttgttacaag 5580
| | | | |
Db 5521 CGTGTGGTTTTATGGAACCAAGTTCCGGGATTTGGATGAACAAGAAATTTGTTTACAG 5580
QY 5581 gagcctgacataacaacactgcagagatatactcaaatggaaggaatttcaagaagaa 5640
| | | | |
Db 5581 GAGCGTGCATTAACCAACTTTCACAGATATCTCACAGATTGGAGGATTTTACGGAGAA 5640
QY 5641 agatttgaagagatgtgttgaaglaatcaagaactctaatccgttagacaagtgttaa 5700
| | | | |
Db 5641 AGATTGGAGAGATGTGTTGAAGTAATCAAAAGCTCTAATCCGTGAGACAAGTGTAA 5700
QY 5701 ttagatccaaagaagcatalatatgaattactctatgtggagccaacttctgacacatat 5760
| | | | |
Db 5701 TTAGATCCTAACAAAGCATATATTGATTAATGATGTGGGCCATATTGTTGACACATAT 5760
QY 5761 gagatgaaggaagaataactactatcgaacaaaatlaacaatctctgctgatacttac 5820
| | | | |
Db 5761 GAGATGAAGGAGAGATACCATTTTGCACAAAATTTCAATCTTCTGCGATTTATGTAC 5820
QY 5821 tgtacaccttacttttagatggcgtgcccattggggaacttcatgaacaaatlaaaaag 5880
| | | | |

```

Db 5821 TGTAACACCTTACTTATGATGCGCGTCCATGCGGACCTTCTGTAACAATTTAAAAAG 5880
Qy 5881 aagaacattctgactacgctcgcacgctcttccctataataaacaaggctcact 5940
Db 5881 AAGACCAATTCAGTACGCTCATGCGCTTCTTATTAATAACAAGGCTCAATGCTCACT 5940
Qy 5941 cataaagaagatcaatccttaacccaattgaagttcctatgtgagacatgcagaaaaag 6000
Db 5941 CATAAAGAAGATCATCTTAACACCAATTAAGTTGATTAAGTGAAGACATGCAAAAAAG 6000
Qy 6001 aacagaagattgagcattgtgacacacatcagatcccccagaccccaaaatgctcagatg 6060
Db 6001 ACACAGAGATTGGCATTTTGCACACATCAGATCCCGACACCCCAAAATGCTTCAAGATG 6060
Qy 6061 gtactccaggagatcgtgtagcaccacagtgatcaaggagccttggagagttgccaggt 6120
Db 6061 GTACTCCAGGAGTCTGTATGGCACCACAGTGAATCAGGGGCTTTGGAAGTTGCCAGGTT 6120
Qy 6121 ttctgtctgaatacctagtgaacccaaagccttcaagacatcataataactgagactc 6180
Db 6121 TTTCTGTGTAATACCTAGTAGACCCAAAGCTCTTCAGACATCAATTAATAACTGCGACTC 6180
Qy 6181 tgccttaagaatttactcaaaaggtgtgaagaatgccttaagaaaaaataagacttaatt 6240
Db 6181 TGTCTTAAGAATTTTACTTAAAGGTTGAAGATGCTTAAAGAAAAAATAGAGCTTAATT 6240
Qy 6241 gggccgggtcacaagaagatcacaagaagaattgggaaacatcttcgccttaagaagac 6300
Db 6241 GGGCCGGTTCACAAAGGATATCAAAAGGGAATTTGGGGAACATATCTCCCTTAAGAAGGC 6300
Qy 6301 cctacagccctagatcacagaaagtcctcagttatcccaagccagattgtctgcccct 6360
Db 6301 CCTACAGCCCTTAGATCACAGAAAGCTCCTCAGTTATCCAACCCAGTATGCTGTCCCT 6360
Qy 6361 gccacaagagattcctctgctgctgaatgagcttgcgcaaatgactcctaactgaatgca 6420
Db 6361 GCCACAGAGATTCTTCACTGATGAGATGACCTTGCAGAAATGATGCTCTAAACTGAATGCA 6420
Qy 6421 ctgtcttatcatcctgcgaagaagccatgatactcaacatcagtgctgaagaatcatgtg 6480
Db 6421 CTGTCTTATCATCTGCAGAAAGGCGCATGTATCAACTGACAGTGAAGATCTAATG 6480
Qy 6481 gaacacaatggaatggaattctggaattatattatcattgaagaatgcagtgccaaag 6540
Db 6481 GAACACAATGGAATGGAATTCGTGAATTTATTATTCATGAAGAATGCGTGGCCAAAG 6540
Qy 6541 aaaataataatgttagattgttaagccttgagaatcatgctatggtttcttaagttcgg 6600
Db 6541 AAAATATCAAAATGTAGATTGTTAAAGCTTGAGAAATCATGCTATGCTTAAATGTTCCG 6600
Qy 6601 gtaacaagctgtatccttctaagaacatttaagctcaagaaggtacactatactacc 6660
Db 6601 GTAACAAGCTGTATCTTTTAAAGACATTTTAATAGCTCAAAAGGTACACTATACCTTTAAC 6660
Qy 6661 attattttacatagctcaaggttaaaaattatattactacttaagttcgtatttttaatt 6720
Db 6661 ATTATTTTACATAGCTCAAGGTTAAAAATTATTCACCTTAAGTTCCTATTTTAAAT 6720
Qy 6721 tatatcacatttagaattcaattttgagacatttaaaagttaagttaagttttta 6780
Db 6721 TATATCACCATTTATAGATTCAATTTTGGACCATTTTAAAGTATGCTATTAATTTAA 6780
Qy 6781 aggtactaaaaaatatgtgaatgttacctgctgcgcagcagcgctc 6828
Db 6781 AGGTACTAAAAAATATGTGAATGTTTACCTGCTGCGCCAGGCGCTC 6828

```

```

RESULT 2
LOCUS AB051558 4311 bp mRNA linear PRI 07-FEB-2001
DEFINITION Homo sapiens mRNA for KIAA1771 protein, partial cds.
ACCESSION AB051558
VERSION AB051558.1 GI:12698086

```

```

KEYWORDS
SOURCE Homo sapiens cDNA to mRNA, clone:pj02581.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Nagase,T., Kikuno,R., Hattori,A., Kondo,Y., Okumura,K. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes.
XIX. The complete sequences of 100 new cDNA clones from brain which
code for large proteins in vitro
JOURNAL DNA Res. 7 (6), 347-355 (2000)
MEDLINE 21082932
REFERENCE 2 (bases 1 to 4311)
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
TITLE Direct Submissio
JOURNAL Submitted (22-Nov-2000) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research, 1532-3, Yama, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
Fax:81-438-52-3914)
FEATURES
source
1..4311
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="pj02581"
/note="vector:BluescriptII SK plus"
1..3910
/gene="KIAA1771"
<1..3910
/gene="KIAA1771"
/note="Start codon is not identified. pj02581 cDNA clone
for KIAA1771 has a 1-bp deletion at the position 498 and a
110-bp insertion after the position 3119 of the sequence
of KIAA1771, respectively."
/codon_start=2
/product="KIAA1771 protein"
/protein_id="BAB21862.1"
/db_xref="GI:12698087"
/translation="KMLEGNHDOHNSLASYIHYVRLPNTYPNSSPGGIGGS
VHYATMARSAVAPASINLNRSLNSNSNDISGPTSPDEVRSLIGSKLDKRSNV
NMGPKPAWGSNPSPSAESNTQAMRSCNRMSHETSFTLOTGLRPLTKLFEEL
ALQWVYSGSVESALQQAFFELMKVSMVHHILYFNKLEPRKSRPERPMDDIA
LVSTIASDVSFRQDTEWERNLNTSLAFNLDDLVSMDRGVESLISCTKYOSKL
VSLPNPSVLSLRDLRLICSHYEVVTLNLSLTPSPASPSVSATYQSQSEFT
NVODOKIAMFELSVFEROQHYLAGVLEIVLAVIDPAEGILGHKKVIMVHNLIS
SHSDPRYSDPQIKARVAMLYPLGLIEMVEYLYDPTFHNGRPICATDYES
EGSNMISQTVNAMAAGTGYLPOLTRGSPILTSFGROHTTSAESSRLILICLWYK
NADETVLOKWFPHDYSYLOLNRLDLILYCVSFETKGVKVERNMSLTFKSKDKRAK
LEBALIGTGARQEWRRSRGQLERSPGSAGFSGENLRKRDTHMRQNTKLDKSR
ABIEHVALIDGNLATEANLIDLTLEIVQVSYVSESKESILGVLKVLHSMCNOS
AVYLHCFFATQALVSKPELLEBEETECADLCURLRHGSSSIGTRSHASSLYL
LKRONEIENFNFAVMOVTNLSLSLVGSONFNEPFRSLKTLITYAEDLELET
TPDOVODIVFNLMILSTPYKREHODPEMLIDLMRIKAGYQTSPLDLTWMQNM
AGKHSRSHHARBAOCIVHASALVAVIYSLMEDRYLYVGCVTTPNISNVLEBSAVS
DDVVSDEBERGICSGFTYESGLVGLLEQAAFSMAGNTEAENYKVLPIHENRD
AKRLSTIHKLQDAFSKIVHSTGWRMGTFRVGFQTRFGDLEDEQVEYKEPAT
KLAIESHRLGEYGERFEDVVEVAKDSVPDKCLDKNKAVIQIYVEPFDYEMK
DRLYEDKNYNLRPMYCTPFLDGRASHELEBOKRKILITTSAPFYITRVVTH
KEEILITPLEVNIEMOKTKOELATRODDPMLMLOVSGSTVVGNGPLEVAO
VFLSETPSPKFLFRHNKRLCTCFKPTKCEBALKKNKSLIPDQKEYORELEKRYHR
LKEALQPLNKRKIPOLYKAVLPVTCRHDSFMSLRKMDL"
BASE COUNT 1312 a 859 c 920 g 1220 t
ORIGIN
Query Match 61.9%; Score 4224; DB 9; Length 4311;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 4276; Conservative 0; Mismatches 25; Indels 7; Gaps 3;
Qy 2508 caaaacttgaaagaaatcatgacacagatgcagaagaagccttcgcatcatatata 2567
Db 1 CAATAACTTGGAAGAAATGATGACCAAGCATGCGACAAACAGCGCTTCTTGATCATATAT 60

```

Qy 2568 tcaatattgtttccgctaccacaatactaccatccatcatccacagctccctg999 2627
|||||
Db 61 TCATATATGTTTCCGCCCTACCAATATCTTACCTTAATTCATCATCCAGGCTCGGGG 120
|||||
Qy 2628 ttgggagagatcaatgcatattatgccaatggctgagatctcggtgagaacctgaaacct 2687
|||||
Db 121 TTTGGAGAGATCATGTCATTAATGCCCAATGGCTAGATTCGGGAGACGACGCAAGCCT 180
|||||
Qy 2688 taatttaaatcgcttcggaagccttagtaataagcaatccagatatatctggaaacctccac 2747
|||||
Db 181 TAATTTAAATCGTTCTCGAAGCCTTAGTATAGCAATCCAGATATATCTGGGAGCTCCAC 240
|||||
Qy 2748 gtccaccaga tga tgaag ttcga tcaatcatccggga ttaagggtttagatcgctccaatlc 2807
241 GTCACCAAGATGATGAGATTGATCATCATCGGAGTAGGGTTAGATCGCTCCAAATTC 300
|||||
Qy 2808 ctgggttaacactggtgtgtccaaaagctgccccatggggatccaaacccagtcacaagtgc 2867
301 CTGGGGTTAACACTGGGTGGTCCAAAAGCTGCCCATGGGGATCCAAACCCAGTCCAAAGTGC 360
|||||
Qy 2868 aagaataacacagagctatgagatcggaagtgtatcgtaatcgtaatgcttgcacacagaaagctc 2927
|||||
Db 361 AGAATACACACAGGCTATGGATCGAAGTTGTATGCTATGCTTGCAACACAGAGACGTC 420
|||||
Qy 2928 aagttcttacaacaacatlaacgaggacgcttacaactaaagccttttcacgagagact 2987
421 AAGTTCTTACAAACGTTATAGGAGCGCTTACCACTAAAAGCTTTTTCACGAGAGACT 480
|||||
Qy 2988 ggccttgcaagtggtgtgttgcagtggaagcttcgggaatcaagctttgcaacaaacctg 3047
481 GGGTTGACAGTGGGTGTTGCAAGTGGCGCTGGGAGTACGCTTTGCAACAGAGCCTG 540
|||||
Qy 3048 gtctcttttgaatlaaagtaagaaagacatgggtgacatcttaacttaatgaataact 3107
541 GTTCTTTTGAATTAATGTAAGCAAGCATGCTGCAACCTTTATCTTTAATGATAAATC 600
|||||
Qy 3108 tgaagctccaagaaagtcgctttccagaagcttcatgaga tgaacatgcaagctctgt 3167
601 TGAGGCTCCACAGGAAAGTCGTTTCCAGAACGTTTCAATGATGATGACGCTTTGT 660
|||||
Qy 3168 cagcagcatgctgtagtatagatagtttcaagatcttcaagaagacacagaatggttgagag 3227
661 CAGCAGATTTGCTAGTATGATATGTTTCAGATTTCAGAAAGCACAGAAATGTTGAGAG 720
|||||
Db 3228 actcaatacaagccttgacatcttcttcaatgactgtgtgtttagttagcagaagagat 3287
721 ACTCAATACAAAGCCTTGGCTTCTTCTCAATATCTGTGTGTATGAGACAGAGATT 780
|||||
Qy 3288 tgtttttagccttataaagtcctgctataaacaagtgcttcaaaagcttactactacc 3347
781 TGTTTTACCTTATTAAGTCTCTGCTATTAACAGGTGTCTTCAAGCTTTTACTCATTTAC 840
|||||
Qy 3348 gaatcccaagtgcttggtggtccttgagctggaatcttctgaatcatctgcagatcaga 3407
841 GAATCCCAAGTGTGTGTGCTTCTTGAAGCTGTGATTTTTCAGAAATCATCTGAGTCATGA 900
|||||
Qy 3408 gcaactatgtataactaaactaacctgcagacttactactcacaactgatatctccatcc 3467
901 GCACATATGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 960
|||||
Qy 3468 ttctgtttcttctgcaacatccagagttctgagtttcttcaagatgtaacagaacaa 3527
961 TTCTGTGTTCTTCTGCAACATCTCAGAGTTTCTGAGATTTCCTAGATGTAACAGACAA 1020
|||||
Qy 3528 gatgccaataatgcttgaatataccggtgcttccgccaacagcaatcttggcagagact 3587
1021 GATTCCAATATATGTTGATTAATTCGTGCTTCCGCCAACAGCATATTTTGGCAGGACT 1080
|||||
Qy 3588 tgggttaacagagctgctgcatctttagaaccttgaatgcttgaagagactgttgatgca 3647
1081 TGTGTTAAACAGAGCTGGCTGTCATTTTGAACCTGATGCTGTAAGAGACTGTTTGGATTGCA 1140
|||||
Qy 3648 taagaagatcatcaaatatgtaacaaattactcaccagtcagactcagaccgcygta 3707

Db 1141 TAAGAAGTATCAATATGATACAAATTTACTTCCAGTCACAGACTCAACCCGGGTA 1200
|||||
Qy 3708 ctctgacctagaataaagctcagatgagcaatgtgtatctactctgtatgtatlat 3767
1201 CTCTGACCCCTCAGATTAAGGCTCGAGTGGCCATGTTATCTACTCTGATTTGATTTAT 1260
|||||
Qy 3768 catggaactgtactcaagctgtatgattttagaagactcacaatcaataaagaagaagacc 3827
1261 CATGGAACCTGTACTCAAGCTGTATGATTTTACGAATCTCACAAATCAACGAGAGAGACC 1320
|||||
Qy 3828 aattgtatagccactgatgaltatgaagaatgaagaggaatgataaagcagaacggt 3887
1321 AATTGTATACCCACTATGATTTATGAAGTGAAGAGGAGATGATGATTAAGCCAGACCGT 1380
|||||
Qy 3888 tgcattgccaatcagcaagagacatcgctccctcaactaaagaagcctgagacttccctct 3947
1381 TGCCATGGCAATCCGACAGGACATCGTCCCTCAACTTAACGAAGCCTGGCAGTTTCCCTCT 1440
|||||
Qy 3948 cagctcaagagatgagcaggaacacactactccttttagcagaatcaagtcgaagcctttc 4007
1441 CAGCTCAACGAGTGGCAGGCAACACTACTCTTTTACGAGAAATCAAGTGAAGCCTTTT 1500
|||||
Qy 4008 gatctgtactatggtgttctcaaaaatgcaaga tgaacagttctcaagaagtggtttac 4067
1501 GATCTGTACTACTTGGGTTCTCAAAAATGCAAGTGAAGAGAGTCTTACAGAGAGGTTTAC 1560
|||||
Qy 4068 agatctctcagcttctgacgttaaacggtctatagatctgcttactctctgtgtcttg 4127
1561 AGATCTCTCAGTCTTGCAGTAAACCGGCTATTAGATCTCTTTATCTGTGTGCTTGG 1620
|||||
Qy 4128 ctgtgatatagaaggaagaaagtggttgaagaaatgaatgagttgacctttagaanaac 4187
1621 CTTTGAATTAAGGAGAAAGGTTTGAACGAATGAATGCTTGACCTTTAAGAAATC 1680
|||||
Qy 4188 aaaaagacatgagagcaaaagcttgaagaagctatcttggagaacatagtgccagagaaga 4247
1681 AAAAGACATGAGACAAAGCTTGAAGAAAGCTTCTTGGAGACATGATGATGCCAGGCAAGA 1740
|||||
Qy 4248 aatgtatcgaggaagccgagagacagctcgagagaagcccatctggaagtgtcttggag 4307
1741 AATGTATCGGCGAAGCCGAGAGACAGCTCGAGAAAGCCATCTGGAAGTGCTTTGGAG 1800
|||||
Qy 4308 tcaagaaatlttagagtgagggagaaagatagactcactggtcctcaaaaacagaagaagt 4367
1801 TCAAGAAATTTGAGGTGGAGAAAGATATGACTCACTGGCTCAAAAACACAGAGAGCT 1860
|||||
Qy 4368 tgaacaaatcaagagcagagatltgaacagagacatgatatgagaaacctgtgctacaga 4427
1861 TGACAAATCAAGAGCAGAGATTGAACACGAAAGCATGATGATGAAACCTGGCTACAGA 1920
|||||
Qy 4428 agcaaaccttaactatlttagatcatagagatgtgtgtctcagaacggttctgtlaacgga 4487
1921 AGCAAACTTAATATATTTAATATATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
|||||
Qy 4488 atccaagaagaagactlttgtagagtgtaaaagtgtaactcaagaalggccgtga 4547
1981 ATCCAAAGAGAGCAATTTGTTGTGTGAGTCTTAAGTGTCTACACAGCATGGCTGTGA 2040
|||||
Qy 4548 ccaaaagtgcagttatctcaacacatgttltgttacaagaagagccttggttccaagtc 4607
2041 CCAAAAGTCAGTTTATCTACACACAGTGTGTTGCTACACAGAGGCTTGGTTTCAAAAGT 2100
|||||
Qy 4608 tccgtgaacttatttgaagaagaagacagagtggtgtgattatgtcctcaggtctct 4667
2101 TCCGAACTCTTATTTTGAAGAAAGACAGAGCTGTGCTGATTTATGCTCAGGCTCTCT 2160
|||||
Qy 4668 ccgaactgtagagtagagctggtgtaataagtgtaacacccagtgctcccttact 4727
2161 CCGAACCTGTAGAGAGTATGCTGTCAATACGCTACACAGCGTACAGGCTCCCTTTACT 2220
|||||
Qy 4728 actaatgagcaaaactttagaatggaaataacttgcgaaggttlaaatgaggtacc 4787

Db 2221 ACTAATGAGCAAACTTGGATTGGGAATACCTTGGCCAGGGTTAAATGACAGTAAAC 2280
QY 4788 aatgtcaatcctcccttggtgggacatctcaaatlttaagaanaattcttaaacg 4847
Db 2281 AATGCACTATCTCTCTGGGGACATCTCAGAAATTTTATGAAGAAATCTTAAAGACG 2340
QY 4848 ttctctaaagcctatattgacatatgtctgaagaagatctggaatctggaggaacaat 4907
Db 2341 TTCTCTAAAGACTATATGACATATGCGAAGAAATCTGSAATTTGAGGGAACAACATTT 2400
QY 4908 tcttgatcaagtlccaagatctggttttcaatctccatatgaatctctctgactgtgaa 4967
Db 2401 TCCATATCAGGTCACAGATCTGGTTTCAATCTCCATATGATCTTCTGTATCTGTGAA 2460
QY 4968 aatgaagaacaaccaaggagatctctgaatgttgatgatataatgaaagaaattggcaa 5027
Db 2461 AATGAAGGAACACCAAGGAGATCCGAAATGTGATGATTAATGTAACAATAATGGCAA 2520
QY 5028 ggtttacagacccctccaga- ---gcgaltgacctggttcagagaacatgagagcaagca 5084
Db 2521 GGGTTACAGACACTCTCCAGATCTGCGATGACCTGCTGCGTTCAGAACTGGCAGCAAGCA 2580
QY 5085 ctcaagaacgaacatcaatctgctgaagctgcacagtgtctagtcacatccagacactgtc 5144
Db 2581 CTCAAGAACGAAACATCTGCTGAAGCTGCACAGTGTCTACTGCACTCAGCAGCACTTGT 2640
QY 5145 tctctaattttgacatgctcgggaagaccggaataatctctctggaatgtgtaactt 5204
Db 2641 TCTCAATATTTTGGACATGCTGGAGACCGGAATATCTCTGTTGGATGTGTAACATT 2700
QY 5205 tcaagaatatctcatctaaatgttttagaagaactctgagtlctcaagatgaatgtgtaactcc 5264
Db 2701 TCAGAAATATTTTCATCTAATGTTTGAAGAAATCTGCGTTCAGATGATGTGATATCTCC 2760
QY 5265 agatgaagaaggatctcgtctcgtgaanaatcttctagtcagacagactgtgggataact 5324
Db 2761 AATGAAAGAAAGTATCTCTCTGGAAATACTTTACTAGTCAAGCACTTGGGGATTACT 2820
QY 5325 ggaagaagagctgcttccctctctatgctgagcaatglatgaagagatlaatgaattta 5384
Db 2821 GGAACACAGAGTCTCTCTCTCTCTATGCTGCGATGTATGAAGAGATTAATGAAGTTTA 2880
QY 5385 caaagtactattctctatctcatgaagctaatctcggaatgcaagaacatctccacaatctca 5444
Db 2881 CAAAGACTTATTTCTAATTCATGAGCTAATCGGATGCAAAACATATCCACAAATTCAC 2940
QY 5445 tggtaactctcaagaagcattcagcaaaaattgttcatcaaggtactgctggggcgat 5504
Db 2941 TGGTAAACTTCAAGAGCAATTCAGCAAAATGTTTCATCAGAGTACTGGCTGGAGCGGAT 3000
QY 5505 gtttggacactattctcgtgtgtgttttlatgaaaccaagtctcgggatttgatgaaaca 5564
Db 3001 GTTTGGACACTATTTTCGTTGGTTGTTTTTATGGAACCAAGTTGGGGAATTTGAGAACCA 3060
QY 5565 agaatttgttcaaaagagctgcaataaacaacttgcagagatatctcacaagattgga 5624
Db 3061 AGAATTTGTTTCAAGAGACCTGCAATAACCAACTTGACAGATATCTCAGACGATTGGA 3120
QY 5625 gggatttttcggagaaagatttggagagatgtgttgaagtaactcaagactcaatcc 5684
Db 3121 GGGATTTTACGGAAGAAAGATTTGAGAGATGTGTTGAAGTAATCAAGACTTAAATCC 3180
QY 5685 tctagaacaagtgtaaatgatctccataacaagcatatattcagattactactgtgagcc 5744
Db 3181 TGTAGCAAGTGTAAATTTGATCTTAACAAGCATATATTCAAGATTACTATGTGAGAGC 3240
QY 5745 ataacttgaacaatltgagaatgaagacgaatcacaactatttcgacaanaaattacaatct 5804
Db 3241 ATACTTTGACACATATGAGATGAAGACAGAAATCACTATTTTCGACAAATAATTAACAATCT 3300
QY 5805 tctgagatcatgattctgcaaccccttactttagatggcgtggccatggggaactca 5864
Db 3301 TCGTCGATTCAATGTACTGTACACCCCTTTACTTTAGATGGCCGCTCCCATGGGAACTTCA 3360

QY 5865 tgaacattcaaaaagaagacacattctgactacgtctcatgcttcttactatataaaac 5924
Db 3361 TGACAAATTCAAAGGAACACCATTTCTGACTACGCTCATGCTTCTCTTATATTTAAAC 3420
QY 5925 aagggtcaatlytcaactcatalaagaagaatcatcttaacacaattggaattgtctatgga 5984
Db 3421 AAGGTCAAATGTCTCACTAAAGAAAGATCAATCTTAAACCAAAATGGAATGTGCAATTTGA 3480
QY 5985 ggaactgcgaagaaagaacagagattgagcatlttgaacacatcagagatcccgagacc 6044
Db 3481 GGAACATGCACAAAAGAACACAGAGATTGGCAATTTGCAACATCCAGATCCGCAAGACC 3540
QY 6045 caaaatgcttcaagatgtactccagagatctglaagcaccaagatgaaatcaggggcctt 6104
Db 3541 CAANAATGCTTCAATGTGTAATCCAGGATCTGTAGGACACACAGTGAATCAGGGCCTTT 3600
QY 6105 ggaagttgcccaggcttctctgtctgaaatcactagtgaccacaagctctcaagatca 6164
Db 3601 GGAAGTTGCCACAGTCTTCTGTGAAATACCTAGTACCCAAACCTCTTCAGACATCA 3660
QY 6165 taataactgagactctgctttaaagatttactaaaggtgtgaaatgaccttaagaa 6224
Db 3661 TAATAACTCGACACTGCTTTAAAGATTTTACTTAAAGGTGGAAGATGCTTAAAGAA 3720
QY 6225 aaataagacttaattgagcggtlcaaaaggatlaaaggaaatg- 'ggaaacta 6282
Db 3721 AAATTAAGCTTAATTTGGCGCGATCAAAAGAGATTCAAAGGAACTGAGAGAACTA 3780
QY 6283 tcttgacctaaagaagccctacagccctagatca- --cagaagctccctcaatltccaa 6340
Db 3781 TCATCGCTTAAAGAGGCCCTTACAGCCACTGCAACAGAAATATCCCTGATTAACAA 3840
QY 6341 gccaagtattgctgtccctgcagcaagagatctcctcaagtctgaatgacttgcgaaat 6400
Db 3841 GGCAGTATTTGCTCTGCTGCTGCGCAGAGATTTCTTCAAGTGAATGACCTTGCAGAAAT 3900
QY 6401 ggaacttcaactgaatgcaatctgttttattatcatctgcaagaagacatglatcaacatc 6460
Db 3901 GGATCTTAACATGAATGACATCTGTTTATTCATCTGCAAGAGCCATGATTAACATTC 3960
QY 6461 gagtgtaaaagatctatcttgaacaacacatggaatcttggaaattatattatc 6520
Db 3961 GAGTGTAAGAAAGATCTTGTGAAACCAATGGAATGGAATTTGAAATTTATTTACT 4020
QY 6521 tgaagaatgcagtgagcaagaataatcaaatgtagattgttlaagcttgagatcatg 6580
Db 4021 TGAAGAAATGCAAGTGGCCAAAGAAATATCAAAATGTAGTTGTTAACCTTGAGAAATCATGG 4080
QY 6581 ctatggttctaaatgtctggglaacaagctgtatcctttaaagacatlttaatgactcaa 6640
Db 4081 CTATGTTTCTTAATGTCTGTGTAACAGCTGTATCTTTTAAAGACATTTTAAATGACTCAA 4140
QY 6641 aggttaactataacttaccattatatacctaagtaaggttaaataattatcaact 6700
Db 4141 AGGTACACTATACATTTTACATTTTATACATAGCTAAAGTAAATAATTTATCACTT 4200
QY 6701 taagtcgtatctttaaattatatacocaattatagatcatcttggaccattttaa 6760
Db 4201 TAACTGCTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 4260
QY 6761 tgaagtaatgcttaatttaaaggtactaaanaalatgtgaatgttac 6808
Db 4261 TGTAGTAACTGTATTTTAAAGTACTATTAATATATGATGATGTTTAC 4308

RESULT 3
AX172880 4144 bp DNA linear PAT 03-JUL-2001
LOCUS AX172880
DEFINITION Sequence 7 from Patent WO0142297.
ACCESSION AX172880
VERSION AX172880.1 GI:14597915
KEYWORDS

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 4144)
TITLE Lu, P., Garman, J. D. and Candia, A. F.
JOURNAL Clasp-3 transmembrane protein
Patent: WO 0142297-A 7 14-JUN-2001;
Arbor Vita Corporation (US)
FEATURES
Location/Qualifiers
1..4144
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Preliminary Clasp-3 cDNA"
CDS
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC43727.1"
/db_xref="GI:14597916"
/translation="RVNRSRLSNSNPDISGTPNPDDEYSIIGSKGLDRNSWNT
GGPKAAPGNSPSPASSTOMDRSCNRMSHETSSFLQTIGRLPTKRLPHEAL
QWVCGSVRESALQOAMPFPELMVKSVMHILYFNDELAPKRSRPFERMDIALY
STIASDIYSRQKDTENWERINTSLAFFLNDLSVMRGVFSLIKSCYQVSSKYS
LPNSVLVSLDELRLICSEHEVVTINLPSLTPPASPSPVSSATSSQSGESTIV
QDOKIANNFELVPRQOHYLAGLVTELAVIIDPDAGEFLGHKVIINMVMHLSH
DSDRYSPQIKARVAMLYPLIGIMEYVQLYDFTETNORRPICIATDYES
GSMISQYVAMAIACTSVOLTRPSGLTSTSGRQHTTFSASRSRLITLIMVAKA
DEVILQKMFNDLSVLOLRILLILYLICVSPPEYKGVFERMNSLTFRKSKDMRAKE
EALIGSGAREMYRRSRGQLERPSSASGSDNLMWRDMHWRONTKLRKRAKE
IEHVALIDGNLATANLILDTLEIVQVTSVESKSLIGVLYKVLHMAQNSQV
YLQHFATQRLVSKPELLEETEOCALICLRLHSCSSIGTILKSHPSASLYLM
RQNFIEINNFARVMQVPMSSLVGTSQNMNEEFLRSKTLITYAEEDLEETFE
PDQVDLVFNHMLISDTYKMKHEOEDPEMLIDLMYIAGYOTSPDLRLTWLONNA
KHSRNSHAEAOCLVSHALVAVEYSMLDRKYLPGVCTRONISNVIEESASD
VSPDEIGICSGKYFTESGLVGLLEQAAASFSMAQNEANYEYKVLPIHEANRAK
KLSTIHGLOBARFKIVHQSITGWERMGTYFRGFTYKFGDDEQEVYKEPALITKL
AEISHLEGFERGERFEDVVEVIKDSNPVCKKIDPKAKYIQTIVPEYVKEPALITKL
ITYDKNYNLRMYCTPFLDGRAGHELHEORFKITLITSHAPEYIKRTAVYTHE
EILTPLEVALIEDMOKTQELAFATODPADPKMLQWLVGSGTGYVNOGPLEAVQV
LSEIPSPKLEFRHNKRLRCLCFDKETKEDALRKNLSLIGPVQKEVORELGLSSP"
BASE COUNT 1252 a 829 c 890 g 1173 t
ORIGIN

Query Match 60.4% Score 4123; DB 6; Length 4144;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4136; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 2693 taatcgtctcgaagccttagtaataagaatccagaatatctggagactccacgtcac 2752
Db 6 TAAATCGTTCGCGAAGCCCTTAGTAATAGCAATCCAGATATATCGGAGCTCCACGTAC 65
QY 2753 cagatgaagaagttcgatcaatcatcgaggagtaagggtttagatcggtccaattccggg 2812
Db 66 CAGATGATGAAGTTGATCATTCATCGGAGTAAGGTTTAGATCGCTCCCAATTCCTGGG 125
QY 2813 ttaacactggtgtgccaaaagctgccccatgaggatccaaccagtcgaagtgcagaat 2872
Db 126 TTAACACTGGTGTGCCAAAAGCTGCCCATGGGATCCAAACCCAGTCAAGTCACAAT 165
QY 2873 caacacagagctatgataagtgatgaatgataatgataatgataatgataatgataatg 2932
Db 186 CAACACAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 245
QY 2933 tcttaacaaca ttaacaggaagcttaccacaataaagccttttccagaggagctgctt 2992
Db 246 TCTTAACAACATTAAACGGAGAGCTTACCACATAAAGCTTTTTCACGAGAGAGTGGCTT 305
QY 2993 tgcagtggtgtgtttgtaagtgagcagcttcggaatcagctttgcaacaagcctgtctc 3052
Db 306 TGCAGTGGTGTGTTCGAGTGGCAGCGTTCGGGAATCAGCTTTGCAACAAGCCTGGTTC 365
QY 3053 ttcttgataataatgtaagaagcattgtgcacacattacttaataataatgtaag 3112

Db 366 TTTTGAATTAATGATGAAGAGCATGGTGACCACTTATTAATGATGAAGCTTGAGG 425
QY 3113 ctccaagaaagatgcttttccagaacgtttccatgataatgataatgataatgataatg 3172
Db 426 CTCGAAGAAAGATGCTTTTCCAGAACGTTTCAATGATGATGATGATGATGATGATGATG 485
QY 3173 cgatgtcgtatgataatgatttcaacgattcaagaagacacagaatgattgagaagctca 3232
Db 486 CGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 545
QY 3233 ataaagccttgatcttcttcaatgataatgataatgataatgataatgataatgataatg 3292
Db 546 ATACAACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 605
QY 3293 ttagcctataaagctcgtatataaagaagtgcttcaaaagccttcaatcctcaatccgac 3352
Db 606 TTACCTTATTAAGTCTGCTGCTATTAACAGAGTGCTTCAAAAGCTTACTATTACGAATC 665
QY 3353 ccagtgcttggtgtccttgaagcctggaatttctcaagaatcactgcaatcagtaagcact 3412
Db 666 CCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 725
QY 3413 atgttaactaaactaacctcagctcagcttacttactcactcagtaatcactcactctctg 3472
Db 726 ATGTTACATTAACCTTACCTGCAAGCTTACTTACTCCTCAGCTGATCTCCTACCTCTTG 785
QY 3473 ttcttctgcaacatcctcagagctctggaatttctcaagaatgataacccaagaatg 3532
Db 786 TTTCTTCTGCAACATCTCAAGATTTCTGAGATTTCTGAGATGATGAACCAAAAGATTG 845
QY 3533 caaatggttgatataatcgttcccttccgccaagaatataattggcagaactgtgt 3592
Db 846 CAATATGTTGATTAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 905
QY 3593 taacagagctgctgtcattttagacccttgatgctgaagagactgtttagattgataaga 3652
Db 906 TAACAGAGCTGGCTGTCATTTTAGACCTGATGCTGATGATGATGATGATGATGATGATG 965
QY 3653 aagtcacatcatatgatacaaatctcctcagtaagtaagtaagtaagtaagtaagtaag 3712
Db 966 AAGTCATCATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1025
QY 3713 accctagaataaagctcgaagtgagtgatgtatcactccttgattgataatcag 3772
Db 1026 ACCCTAGATTAAGCTTCGAGTGAGCTGATGATGATGATGATGATGATGATGATGATGATG 1085
QY 3773 aaactgtacctcagctgataatlttaccagaacatcaacaatcaagaagaaagcaaat 3832
Db 1086 AAACGTGACCTCAGCTGATGATTTTACAGAAACCTCAACATCAGAGAGAAACCAATTT 1145
QY 3833 gataagcactgtgatttgaagaatgagcgggaagtatgataagccaagcgtttgcca 3892
Db 1146 GTATACCACTGATGATTTGAAAGTGAAGCGGAAGTATGATGATGATGATGATGATGATG 1205
QY 3893 tggcaatcgcaaggaatcgtgctcccaactaacaagagcttgagcttccctcccaagt 3952
Db 1206 TGGCAATCGCAGGACATGCTGCTCCCACTAACAAGGCTTGCGAGTTTCTCTCCACAGT 1265
QY 3953 caagagtgagcagcaacaacactacctttcagcagaatcaagtcgaagcctttgact 4012
Db 1266 CAACGAGTGGCAGGCAACACTACTCTTTCAGCAAGAAATCAAGTGAAGCCTTTGATGAT 1325
QY 4013 gttactttggtgtctcaaaaatgcaatgaaagcttcaagaagtggtttacagatc 4072
Db 1326 GTCTACTTGTGGTGTTCAAAATGAGATGAAGAAAGTCTTACAGAAAGTGATTAAGATC 1385
QY 4073 tctcaatcttgcaagctaaacccgctatagatccttcttactctgtgtgtgtgtgtgtgt 4132
Db 1386 TCTCAATCTTGCAGCTAAACCGGCTATTAGATCTCTTATCTCTGTGTGTGTGTGTGTGTGT 1445
QY 4133 agtataaaggaagaaagtgatttgaacgaatgaatgacttgaacttgaagaatcaaaag 4192

Dh 1446 AGTATAAGGAAAAAGGTGTTGAACGAATGANTAGCTTTAGACCTTTAAGAAATCAAAAG 1505
Qy 4193 acatgagagcaaaagcttgaagaagactatctcttggaacataagtgccaggaagaatg 4252
Dh 1506 ACATGAGGCAAAAGCTTGAANAAGCTATTCTTGGAGCATATAGTGCCAGGCAAGAAATGG 1565
Qy 4253 taaggcgaagccgaagcaagctcgagaagaagcccatctggaagtgccttggaaagtc 4312
Dh 1566 TACGGCGAAGCCGAGGACAGCTCAGAGAAAGCCCATCTGGAAGGCTTTGGAAGTCAAG 1625
Qy 4313 aaaattgaggttggaagaagatatgactactggtgcgaagaacagaagaagcttgaca 4372
Dh 1626 AAAATTGAGGTGAGGAAGAAATATGACTCATGGCGCTCAAAAACAGAAAGCTTGACA 1685
Qy 4373 aataagagcagagattgaaacagcaagcaactgattga tggaaacctggccacagaagcaa 4432
Dh 1686 AATCAAGAGCGAGATTTGAACACGACACTGATTGATGGAAACCTGGCTACAGAAAGCAA 1745
Qy 4433 acctaatcatttagatacatatagagatgltgtgttcagacggttctctgtaacggaatcca 4492
Dh 1746 ACCTAATCATTTTAGATATACATTAGATGATGTTGTTAGACCGTTCGTAAACGGAATCCA 1805
Qy 4493 aagagcagattctctgtgtgaggtgtctaaagtgctactacaagaatgctgttacaacaa 4552
Dh 1806 AAGAGAGCAATTTGTGGAGTGTCTAAAGTGTACTACACAGCATGGCCTGTAACCCAAA 1865
Qy 4553 gtgcagttatctacaacactgtttgtctacagagagccttggttcaagaattctctg 4612
Dh 1866 GTGCAGTTATCTACAAACACTGTTTGTCTACACAGAGACCTTGTTCAAAGTTTCTCG 1925
Qy 4613 aactcttatttgaagaagagacagagcagtgctgattatgacctaaaggtctctccgac 4672
Dh 1926 AACCTTTATTTGAAGAAGAGACAGACAGAGTGTGATTTATGCTCAGCGCTTCCTCCGAC 1985
Qy 4673 actitagaagagatgagtgaacaataagtcacagcccaagtgctcccttactactaa 4732
Dh 1986 ACTGAGAGTAGACATGGATACAAATACGATCACCCAGGCTCCCTTACTACTATAA 2045
Qy 4733 tgaagcaaaacttggagatggagataacttgcagagttaaagtcaagttacaaatgt 4792
Dh 2046 TGAAGCAAAACTTTGAGATTGGGAATTAACTTTGCCAGGGTTAAATGCAGGTACCAATGT 2105
Qy 4793 caactactcctctgtgtggacatctcagaatttaatagaagaattcttaagcgtctcc 4852
Dh 2106 CACATATCCTCTGTGGTGGCCACATCTCAGAAATTTAATGAAGATCTTAAAGAGCTTCTC 2165
Qy 4853 taaagacatattgacatatgtctgaagaagatctggaattgagggaaacaacattctctg 4912
Dh 2166 TAAAGACTATATTTGACATATGCTGAAGAAGATGTGAATTGAGGGAACAACATTTCTCG 2225
Qy 4913 atcaggtcagaatcgtgttctcaatctccatgattcttctgactactgtgaatga 4972
Dh 2226 ATCAGGTCCAGGATCTGGTTTCAATCTCCATATGATTTCTTTCGATACGTGTAATAATGA 2285
Qy 4973 aggaacaccaggaagatctgaaatgltgattgatactaatgtacagaattgccaaggtt 5032
Dh 2286 AGGAACACACAGAGAGATCCTGAATGTTGATTGATTAATGTACAGAAATTTGCCAAGGTT 2345
Qy 5033 accagacctcccaaga---gagattgacctgtgttcagagaatgagcaggaagcaccag 5089
Dh 2346 ACCAGACCTCTCCAGATCTGCAATGACCTGGTTGCAAGAAATGGCAGGCAAGACATCAG 2405
Qy 5090 aacgaagaatcatgtctgaagctgcacagtgctagtcaccactgaagcaactgttctg 5149
Dh 2406 AACGAGCAATCATGCTGAACACTGCACAGTGTAGTCCATCAGCAGCACTTGTTCGTG 2465
Qy 5150 aatatttgaagatgctggaagaccggaataatcttccbtgtggatgtgtlaaacatttca 5209
Dh 2466 AATATTGTAGCATGCTGAGGAGACCGGAAATATCTTCCGTGGAGTGTAAACATTTTCA 2255
Qy 5210 atattctcaatgatttgaagaagatctggtctcagatgagtgttatctccaatg 5269
Dh 2526 ATATTTCATTAATGTTTAAAGAAATCTGGGTTCTAGATGATGTGTATCTCCAGATG 2585

Qy 5270 aagaagtaactgtctcttggaataatacttactgaagtcaggacttgggattactggaac 5329
Dh 2586 AAGAAGTATCTGCTGTGAAAATACTTTACTGATCTCAGGACTTGTGGATTACTGGAAC 2645
Qy 5330 aagcagctgtctctctctatgtctgcatgatagaagcagttaaatgaagttacaag 5389
Dh 2646 AAGCAGCTCTTCTCTCTTATGCTGGCATGTATGAAGCAGTTAATGATTTTCAAG 2705
Qy 5390 tacttaactctctatgaagctaagtggaatgaagaataactcaacttaactgta 5449
Dh 2706 TACTTATCTTCTATCTGAAGCTAATCGGATGCAAGAAACTATCCAAATTCATGGA 2765
Qy 5450 aactlcaagaagcatlcaacaaatgttctcatcaagactgtcgtcgggaagcgaatgtt 5509
Dh 2766 AACTTCAGAAGATTCACCAAAATTTGTCATCAGAGTACTGCGGGGCGGAGTTTG 2825
Qy 5510 gcaactattctgtgtgttltatlgaaaccaagltcggggaatttgatgaagaagat 5569
Dh 2826 GCACCTATTTTGTGTTGTTTATGAAACCAAGTTGCGGATTTGGATGMAACAAGAT 2885
Qy 5570 tigtlttaagaagagcctgtaataaaccactgcaagagatctcaagattgaggaat 5629
Dh 2886 TTGTTTCAAGGAGCCTGCAATTAACCAACTTGACAGATATCTCACAGATTGGAGGAT 2945
Qy 5630 tttaacgagaagatttggagagagatgtgttgaagtaatcaagaacttaactctgtag 5689
Dh 2946 TTTCAGGAGAAAGATTTGGAGAGAGATGTTGAAGTATCAAAACACTATTCCTGTAG 3005
Qy 5690 acaagtgtaatagatctcaacaaagcatatatacgaattacactatgtgagccatct 5749
Dh 3006 ACAAGTGTAAATTAGATCTTAACAAGGCATATATTCAGATTACCTATGTGGAGCATCT 3065
Qy 5750 ttgacacataagatgaaagacaagatacactatttgcagaagaataataactctgtc 5809
Dh 3066 TTGACACATATGATGAAAGACAGATACCTATTTGACAAAATTTCAATCTTCGTC 3125
Qy 5810 gattcatgactgtacacacttactttagatgagccgttcccatggggaacttcatlgaa 5869
Dh 3126 GATTCAATGATGTACACCTTTACTTTAGATGGCGGTGCCATGGGGAACCTTCATGAA 3185
Qy 5870 aattcaaaaggaagacacttctgaactggtctcaatgacttccctatataaacaagg 5929
Dh 3186 AATTCAAAAAGGAAGACCAATCTGACATCTCATGSCCTTCTTATTAATAAACAGGG 3245
Qy 5930 tcaatgactcaataagaagagatcatcttaaccocaaatgaagtgtctatgtggaca 5989
Dh 3246 TCAATGTCACTTAATAAGAAAGATCATCTTAACCAATTAAGTTGCTATTGAGGACA 3305
Qy 5990 tgcagaaaaagacacagaggttggcaattgcaaacacatcaagatcccgacagaccacaa 6049
Dh 3306 TGCAGAAAAAGACAGAGAGTGGCATTTTGCAAAACATGAGATCCCGAGACCCCAAAA 3365
Qy 6050 tgcctcaagatgtaactcaaggaatctgtaggacaacagtgaaatcaagggtccttggaa 6109
Dh 3366 TGCTTCAGATGTGATCCAGGGATCTGTAGGCCACCACTGATATCAGGGGCTTTGGAA 3425
Qy 6110 ttgcccaggttttctgtctgtaaatcatcagtgagcccaagctctcagacactaata 6169
Dh 3426 TTGCCCAAGGTTTCTGTGAAATCTTAGTGACCCAAAGCTCTTCACATCAATTAATA 3485
Qy 6170 aacttgacactgtccttlaagaatttactcaaaagtgtaagaatgccttlaagaanaata 6229
Dh 3486 AACTGGACATCTGCTTTAAAGATTTTACATAAAGGTGTAACATGCTTTAAGAAAAATA 3545
Qy 6230 agagcttaatttggccggttcaaaagagatcaaaaggaatttggggaactatctctgc 6289
Dh 3546 AGAGCTTTAATTGGCCGCGTTCAAAAAGAGTATCAAAAGGAATTTGGGAAACTATCTGC 3605
Qy 6290 cttaagaagggccctacagccctagatcacagaagcctctcagttatccaagcagatct 6349
Dh 3606 CTTAAAGAGGCCCTTACAGCCCTTAGATCACAGAAAGTCCCTCAGTTATCCAACCCAGTAT 3665

QY	6350	gcttctcccttcacacagagatccttcagtcgtgaatgagcttcgcacaaatgagatccta	6409
Db	3666	gcttctcccttcacacagagatccttcagtcgtgaatgagcttcgcacaaatgagatccta	3725
QY	6410	aactgaatgcactctgtttatcatcctgcgaagaagccatgataccaactgagtgta	6469
Db	3726	aactgaatgcactctgtttatcatcctgcgaagaagccatgataccaactgagtgta	3785
QY	6470	aagatcctatttggaaccacaatcagaaatggaattcctcgaaatattatcattgaaagatg	6529
Db	3786	aagatcctatttggaaccacaatcagaaatggaattcctcgaaatattatcattgaaagatg	3845
QY	6530	cagtgcccaagaanaatcacaaatgtagatctgtttaaagccttggagaatcaagctatggttt	6589
Db	3846	cagtgcccaagaanaatcacaaatgtagatctgtttaaagccttggagaatcaagctatggttt	3905
QY	6590	ctaattgctcggtaacaaagctgtatcctttaaagacatttaagtactcaaaagctact	6649
Db	3906	ctaattgctcggtaacaaagctgtatcctttaaagacatttaagtactcaaaagctact	3965
QY	6650	atacgtttacacatttatatacattagcgtcaaggttaaataattatcacttaagtctgt	6709
Db	3966	atacgtttacacatttatatacattagcgtcaaggttaaataattatcacttaagtctgt	4025
QY	6710	attctttaaattatatacaccatttataagatcaattcttggaccacattttaaagttagtaat	6769
Db	4026	attctttaaattatatacaccatttataagatcaattcttggaccacattttaaagttagtaat	4085
QY	6770	gctattttttaaagctactcaaaaataatgtaagtcttacctcgtgcgcgcagagctc	6828
Db	4086	gctattttttaaagctactcaaaaataatgtaagtcttacctcgtgcgcgcagagctc	4144

RESULT	4				
LOCUS	AX172932	4144 bp	DNA	linear	PAT 03-JUL-2001
DEFINITION	Sequence 59 from Patent WO0142297.				
ACCESSION	AX172932				
VERSION	AX172932.1	GI:14597917			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 4144)				
AUTHORS	Lu, P., Garman, J. D. and Candia, A. F.				
TITLE	Clasp-3 transmembrane protein				
JOURNAL	Patent: WO 0142297-A 59 14-JUN-2001; Arbor Vita Corporation (US)				
FEATURES	Location/Qualifiers				
source	1..4144				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/note="human CLASP-3 cDNA"				
	197..3610				
	/note="human CLASP-3"				
	/codon_start=1				
	/protein_id="CAC43728.1"				
	/db_xref="GI:14597918"				
	/translation="MDRSNRMSSHTETSPFLQTLGRLPKKLFHEELALQWVCSG SVRSALQQAWEFFELWVSKVHHLYTNDKLEARKSRFERPMDDIALVSTIASD VSRQAKTEWERTENLTSIAFLINDLVMQGEVFSIKSCYKSSKSLVLPBVSIV LSLDLRLTICSHHEHYVTNLPCSLTLPASPEPSATSSQSGSTINVDOKIANM FELSVPEROOHYLAGLVITPELAVITLDPDGGFLGKHYINMHNILSSHDSPRSV DPQKARVAMLYLEPLIGTIMETVYQLDPTFNHNRGPICIAIDVDESEGSMTSQT VAKIAGTSVQLTRPGSEFLTSTSGQHFTFSSESSRSILCLLWALKNADETVLQR WFTDLVSLQNLRLDLILYLCVSCREYKGRKVEEMNSLTFKSKDMRAKLEALIGSI GARDEMVRSGQLERSPSGSAGFSGENLRKRMKMTIMRWOTETKLDGSRALIEHALI DGNLTATANILITDLTEIVVOTVSVETSKSIIIGGLVKVILSHMACNOMAYVLOHCR TORNLATKPEPLLEPTEECADCLILRLRCSSIGTIRHSPASLYLLMRQNETIG NNRARVMQYPMSSLISLVGTISQNEEFLRSRLTILTYAEEDLEKETFPDGVQDI VFNLHMLTSDTVKKRHEQDEPMLIDILRYAKCYQSPDLRLYLNQMGKHSERSN HAEKAOCLVSAALVAEYLSMLTERKTLPGVCYFQVQNSISVLEESAASDVVYPPDEE				
CDS					

BASE COUNT	1252	a	829	c	890	g	1173	t	ORIGIN
Query Match	Best Local Similarity	60.4%	Score 4123;	DB 6;	Length 4144;				
Matches 4136;	Conservative	99.9%;	Pred. No. 0;	Mismatches	0;	Indels	3;	Gaps	1;
QY	2693	taaatcgcttcgaagccttagtaataagcaatccagatatatctggaatcccaagtcac	2752						
DB	6	TAAATCGTTCTTCGAAAGCCCTTAGTATACCAATCCAGATATATCGGAGCTCCACGTCAC	65						
QY	2753	cagatgataaagcttcgcatacatccgggaataagagtttagatcgcctccaatccctggg	2812						
DB	66	CAGATGATGAATTCGTGATCAATCATCGGGAGTAAGGGTTTAGATGCGCTCCCAATTCGGG	125						
QY	2813	ttaaacactggttgatccaaaagctggcccaatggggatccaaaccccaagtcagaatg	2872						
DB	126	TTAAACACTGGTGGCTCCAAAAGGTGCCATGGGGATCCAAACCCAGTCCAAAGTCAGAAAT	185						
QY	2873	caaacacagctatgatactgaagttgtaactgatagtcttcgcacacagagaagctcaagt	2932						
DB	186	CAACACAGGCTATGATGATGGAAGTTGTAATGCTATGCTTCCACACACAGACGCTCAAGTT	245						
QY	2933	tcttcaacaacatlaacaggagcagctaccactaaagctcttccacagaagagcttgct	2992						
DB	246	TCTTCAACAACATTAACGGGACGCTTACCACTAAAAACCTTTTCCACAGAGAGCTGGCTT	305						
QY	2993	tgcagttggttgcttgcaagttgcagcgcttcgggaatcaagctctgcaacaagcctggtctc	3052						
DB	306	TGCAGTGGGTTGTTTGCAGTGGACCGCTTCCGGAAATCAGCTTTCACAAACCCGTGTTCT	365						
QY	3053	tttttgaattatgataagagcagctggtgacacatttactttaaataaacttaagg	3112						
DB	366	TTTTGAATTAATGTTAAAGAGCATGTCGCCATTTTATCTTAAATGATTAACCTTGAGG	425						
QY	3113	ctccaaggaanaagtcgtttccagaagctttcatgataagacatgacatgacagctctgtcaga	3172						
DB	426	CTCCAAAGGAAAAAGTCGTTTCCAGAACGCTTTCATGAGATGACATGTGACGCTCTTGTCAACA	485						
QY	3173	cgatgacagatgataatgttcacagatttcgaagagcacagaataagtttgaagagactca	3232						
DB	486	CGATGCTAGATGATTAATGTTTCCAGATTTCAGAAAGGACACAGAAATGGTTGACGACTCA	545						
QY	3233	atacaagccttgcatctcttcataatgatactgttgcgtatgataagacagaagatttgtt	3292						
DB	546	ATACAAAGCCTTGCAATCTTTTCTCATATGATGTCGTGTTATGAGACAGAGATTTGTGTT	605						
QY	3293	ttagacctataaagtctcgcgtataaaacaggtgtcttcaaaagcttactcatcaccgaatc	3352						
DB	606	TTAGCCTTATATAAGCTCCGTATATAACAGGGTCTTTCAAAACCTTACTCATTAACCGAATC	665						
QY	3353	coaggtcttcgtgtccctctgaagcctggaatttcttaagaaatcatctgcagatgaagcact	3412						
DB	666	CCAGGTTCTGCTGTCCTTGAGGCTGGATTTTCTACGAATATCTGCAAGTCATGAGCACT	725						
QY	3413	atgtataacttaacttaccctgcagcttacttacttaccacctgcatactcatcactcttg	3472						
DB	726	ATGTACTATTTAAACCTTACCTCGACAGCTTACTTACTCCACCGCAATCCATCACTCTTG	785						
QY	3473	tttctctgcgaactctcaaggtctcgtatatttctaagaatgatacaagacaaaagaattg	3532						
DB	786	TTTCTTCTGCAACATCTCAGAGTTCTGGATTTTCTACGAATGATCAAGACCCAAAAGATTG	845						
QY	3533	caaatatgttgaatlaacgcgtgctctccgcgaagacatlaatttggcaggaacttgt	3592						
DB	846	CAAAATATTTGAATTAATCGGTGCTTTCCGCAACAGCAATTAATTTGGCAGGAACTTGCT	905						

QY	3593	taacagagctgctgctgcatcatttagaccctgatagtcgaagcctgtttagatltgataaga	3652
Db	906	TAACAGAGCTGCGTGTTCATTTTAGACCCGTAGTCGAAGGAGCTGTTGGATTGCATAGA	965
QY	3653	aagcatcaatatgftaaccaattactctcaagfcaagactcaagaccgcygactctg	3712
Db	966	AAGCATCAATATGFTAACCAATTACTCTCAAGFCAGACTCAAGACCCGGGATACCTG	1025
QY	3713	accctcagataaagctcgaagtgcgaatctgtatcatcctctgattgtaatactg	3772
Db	1026	ACCCCTCAGATTAAGGCTCGAGTGGCCATGTGTATCTACCTGTATGTGTATATCANG	1085
QY	3773	aaactgtacctcagctgtatgatatlttaagaactcacaatcaacaggaagaccaatlt	3832
Db	1086	AAACTGTACCTCAGCTGTATGATTATTCAGAAACTCACAAATCACAGGAGAACCAATTT	1145
QY	3833	gtataagccactgattgattatgaagaagfagaagggaaatgatatgaagccagacgtttcca	3892
Db	1146	GTATAGCCACTGATGATTATGAAGAAGFAGAAGGGAAATGATGATAAGCCAGACCGTTGCCA	1205
QY	3893	tggaatcgcaggaagacatcgtgtccctcaactaacaaggcctgcagtttccctccacgt	3952
Db	1206	TGGCAATCGCAGGGACATGTCGTCCCTCACTAACAAAGGCTTGGAGTTTCCCTCCACGT	1265
QY	3953	caacgagttgcgcagggacaacatacctcttccagcagaatcaagtcgaagcctttgactc	4012
Db	1266	CAACGAGTTCGCAGGCAACACACTACTTTTCAGCAGAAATCAAGTCGAAAGCCTTTGGATCT	1325
QY	4013	gtcactcttggtttccaaaatactgacatgtaaacagttcctaagaagfgygtttcaagatc	4072
Db	1326	GTCACTTTGGGTTCTCAAAAATGCGATGAAACAGTTCTACAGAAAGTGGTTTACAGATC	1385
QY	4073	lctcagctttgcagcttaaacgcgtctatagatctgcttatctcgtgtgtctgtcttg	4132
Db	1386	TCCTCAGTCTTGACACTTAACCGCGCTATFAGATCTGCTTATCTGTGTGTCCTGCTTGG	1445
QY	4133	agataaagggaaaaaaggttlttgaaacgaatgaatagctcttgaccctttaagaatccaag	4192
Db	1446	AGTATAAAGGGAAAAAGGTTTTGAACGAAATGATATGCTTGACCTTTAAGAAATCCAAG	1505
QY	4193	acatgagagcaaaactctgaaagaactctctcttgagagcataagtcgaagcaagaatag	4252
Db	1506	ACATGAGAGCAAACTTGAAGAAAGCTATTCTTGGAGCATATAGTCCCGACGACAGAAATGG	1565
QY	4253	taaggcgaagccgagagcaagctcgaagagaaagcccatctgyaagtgccctltggaagtcag	4312
Db	1566	TACGGCGAAGCCGAGGACACTCGAGAGAAAGCCCATCTGGAAGTGCCTTTGGAAGTCAAG	1625
QY	4313	aaaatttgaggttgaagaaagaatatgaatccactctgctcgtcaaaaacagaagaagcttgaca	4372
Db	1626	AAAAATTGAGGTGAGGAAAGATATGATCACTACGGCGTCAAAAACAGAGAAAGCTTGACA	1685
QY	4373	aatcaagagcagagatatggaacacgaagacatgtatgataaagcctggcttaagaagaa	4432
Db	1686	AATCAAGCGAGAAATTGGAACACGAAACACTATGTTATGGAAACTGGCTACAGAAAGAA	1745
QY	4433	acctaatcattttagatacatitgagatgtgtgtltaagacglttctgtlaacgyaatcca	4492
Db	1746	ACCTATATCATTTTGTAGATACATFAGAGATTGTGTGCAGACCGTTTCTGTAAGCAATCCA	1805
QY	4493	aagagagatctctgtgtgtgtgtctaaagtgctactatacaacgcgatgcgtgtataccaa	4552
Db	1806	AAGAGAGCATTTCTGTGTGTGTCTAAAAATGCTACTTACACAGCATAGGCGCTGTATAACCAA	1865
QY	4553	gtgcagattactatacaacactgtttgtctcaacagagagccttggtttcaaaagtltctg	4612
Db	1866	GTGCAGTTTATCTACAAACACTGTTTGTCTACACAGAGAGCCTTGTGTTCAAGTTTCTGTG	1925
QY	4613	aactctatttgaagaagagacagagcagtgigtgtatltatgtcctcaaggttctcgac	4672
Db	1926	AACCTTTTTTTGAAGAAGAGACAGAGCAGTGTCTGTATTATGCTTCAGAGCTTCTCCGAC	1985

QY	4673	actbaacagatagatcgtgfracaatacggtaacacccagtcgccccttaactactaciaa	4732
Db	1986	actgtagagagatgagctgactacatcaggccacacccagtcgccccttaactactaciaa	2045
QY	4733	tgagcaaaactcttgagatttggaataaactcttgccaggtttlaaaatgcaggtlacaatgct	4792
Db	2046	tgagcaaaactcttgagatttggaataaactcttgccaggtttlaaaatgcaggtlacaatgct	2105
QY	4793	caatactcctctggtgtagacatctccagaatttaagaagaattcttaagaagcttctc	4852
Db	2106	cacatcctcctctggtgtagacatctccagaatttaagaagaattcttaagaagcttctc	2165
QY	4853	taagaacatacttgcatactgcgcgaagaagatctgaaatttgagggaacaacactctcgt	4912
Db	2166	taagaacatacttgcatactgcgcgaagaagatctgaaatttgagggaacaacactctcgt	2225
QY	4913	atcagttccaggaatctggttttcaatctccatagtatcttctgcatactgtaaaatga	4972
Db	2226	atcagttccaggaatctggttttcaatctccatagtatcttctgcatactgtaaaatga	2285
QY	4973	aggaacccaggaagatccctgaatctgtaattgataatgaagaatttgcaaggtt	5032
Db	2286	aggaacccaggaagatccctgaatctgtaattgataatgaagaatttgcaaggtt	2345
QY	5033	accagacctctccaga---cgatctgacctggtctgcagacaatggcaggaacagactag	5089
Db	2346	accagacctctccaga---cgatctgacctggtctgcagacaatggcaggaacagactag	2405
QY	5090	aacgaagaatcatctgctggaagctgcacagtgtctagtccaatccagaacactgttctgt	5149
Db	2406	aacgaagaatcatctggaagctgcacagtgtctagtccaatccagaacactgttctgt	2465
QY	5150	aatatctgagatctgtaggaacccgaatatctctcgttggaatgctgaacattccga	5209
Db	2466	aatatctgagatctgtaggaacccgaatatctctcgttggaatgctgaacattccga	2525
QY	5210	atattctcaatgattttagaagaatctgcggtctccagaatgattgtagtctccagatg	5268
Db	2526	atattctcaatgattttagaagaatctgcggtctccagaatgattgtagtctccagatg	2585
QY	5270	aagaagratctgctctcggaataactcttgcatactgcaggaactctggatcttaccgagac	5322
Db	2586	aagaagratctgctctcggaataactcttgcatactgcaggaactctggatcttaccgagac	2645
QY	5330	aagcagctgcctctctctcctatctgctgcgaatgataagaagattlaaagaatttacaag	5388
Db	2646	aagcagctgcctctctctcctatctgctgcgaatgataagaagattlaaagaatttacaag	2705
QY	5390	taactatctcctatctcagaatctcaatccggatgcgaagaagaactatccacaattctgta	5449
Db	2706	taactatctcctatctcagaatctcaatccggatgcgaagaagaactatccacaattctgta	2765
QY	5450	aacttcaagaagaatctgagcaaaattgttcatacagatctgctctgagagcgatgttg	5500
Db	2766	aacttcaagaagaatctgagcaaaattgttcatacagatctgctctgagagcgatgttg	2822
QY	5510	gcacatattctcgigtgtgttttataagaaccaagtctcggygaatttgatagaacaagaat	5565
Db	2826	gcacatattctcgigtgtgttttataagaaccaagtctcggygaatttgatagaacaagaat	2885
QY	5570	tgatttacaaggagctgcgcaataaccaaattgcagagatattctcaaatgtggaggt	5622
Db	2886	tgatttacaaggagctgcgcaataaccaaattgcagagatattctcaaatgtggaggt	2945
QY	5630	tttaacggagaagaatttgagagagatgttggttgaagttaacaagaacttaactctgtag	5685
Db	2946	tttaacggagaagaatttgagagagatgttggttgaagttaacaagaacttaactctgtag	3005
QY	5690	acaagtctgaatagatccctacaagaagcatataatcagaattactctatgtgagccaatct	5749
Db	3006	acaagtctgaatagatccctacaagaagcatataatcagaattactctatgtgagccaatct	3065
QY	5750	ttgacaacataagaatgaagaagcaatcccaatttcacagaaaattacaactctctctc	5809

```
|||||
Db 3066 TTGACACATATGATGAGAGACAGATACCTATTGGACAAAATATACATCTTCGTC 3125
QY 5810 gatctatgactgcttaaccccttactttagatgcccgcgcccacgaggaaactcaltgaac 5869
Db 3126 GATTCATGTAGCTACACCTTACTTTAGATGGCCGTCGACATGCGGAACTTCATGTAAC 3185
QY 5870 aattcaaaaggaagaccattctgactagctgctcatgaccttctctatatttaaaagaagg 5929
Db 3186 AATTCAAAAGGAAGACCATTTGACTAGCTCTCATGCTCTTCTTATATTAAACAAAGG 3245
QY 5930 tcaatgctacataaagaagagatcatcttaacccaattgaagtgtgctattgagaca 5989
Db 3246 TCAATGCTACTCATTAAGAGAGATCATCTTAACACCAATTGAAGTCTATTGAGACA 3305
QY 5990 tgcagaaaaagacagagaggttgcatlttgacacatcagatcccgacaccccaaaa 6049
Db 3306 TGCAGAAAAAGACACAGAGGTTGGCAATTGGCAACATCAGATCCGCGACCCCAAAA 3365
QY 6050 tgcctcagatggttactccagagatctgtagacacacagtgaaatcaggggcctttggaag 6109
Db 3366 TGCCTCAGATGTTACTCCAGGGATCTGTAGGCAACACATGATAGGGGCTTTGGAAAG 3425
QY 6110 ttgcccagggttctctgctcgaatacactagtgacccaaggctctcagacataata 6169
Db 3426 TTGCCAGGCTTTTCTGCTGTAATACCTAGTACCCAAAGCTTTTCAGACATCATATA 3485
QY 6170 aactcgcactgctctttaaagatttactaaagggtgtaagagtgcttaagaaaaata 6229
Db 3486 AACTGGAAGCTGCTTTAAAGATTTTACTTAAAGGTGTGAAGTCCCTTAAGAAAAATA 3545
QY 6230 agagcttaatttgccgcggttccaagaagagatcaagaaggaattggggaaactatctgc 6289
Db 3546 AGAGCTTAATTTGGCGGTTCAAAAGGATATCAAGGGAATTTGGGAAACTATCTTGGC 3605
QY 6290 cttaagaagagccctcagccctagatccaagaagctccctaagtatccaagccaagatt 6349
Db 3606 CTTAAGAAGGCCCTCAGGCCCTAGATCAAGAAAGTCCCTCAGTTATCCAAAGCCAGTAT 3665
QY 6350 gcttgcacctgcacagagattcctcagtcgaatgagagcttgcgaatggatctcta 6409
Db 3666 GCTTGTCCCTGCGCACAGAGATCTTCAAGTCGAATGAGCTTTGCCAAAAGAGATCTCTA 3735
QY 6410 aactgaaatgcaactgtttatcatctgcgaagagcatalatcaaatcgaagtgtaa 6469
Db 3726 AACTGAATGCACTGTGTTTATTCATCTGCAAAAGACCATGTATTCACATCGAGTGGA 3785
QY 6470 aagatctattgaaacccaacatggaatggaattctggaattatattcatctgaagaatg 6529
Db 3786 AAGATCTATTTGGAAACCAACATGGAATGGAATCTGGAAATTAATTCATTGAAGATG 3845
QY 6530 cagtgccaagaanaatacatcaatgtagatgtttaaagcttgaagatctgcttggtt 6589
Db 3846 CAGTGGCCCAAGAAATATCAAAATGTAGATTGTTAACGCTTGAAGATCAATGAGCTAAGGTTT 3905
QY 6590 ctaatgttcgggttaacaagctgttactctttaaagacattttaaagactcaaaaggtacact 6649
Db 3906 CTAATGTTGCGGTACAGAGCTGTATCTTTAAAGCATTTTAAGACTCAAAAGGTACACT 3965
QY 6650 ataacttaccatttataaccatagctaaaggtttaaatttcaacttgaagtcgt 6709
Db 3966 ATACATTTACCATTAATTAATACATAGCTAAGGTTAAATAATTTATTCACCTTAAGTTCGT 4025
QY 6710 atttttaattatataccattatagatcatcttgagaccattttaaattgaagaaat 6769
Db 4026 ATTTTATTAATTAATATCCCATTTATATGATTCATTTTGGACCAATTTAAATGATGATAT 4085
QY 6770 gctattttaaaggtactaaaaaataatgtaagtgttcaacctgctgcgcgaaggctc 6828
Db 4086 GCTTATTTTAAAGGTACTAAAAAATATGTGAATGTTTACCTCGAGCGGCGGACAGGCGCTC 4144
```

RESULT 5

```
AX172980
LOCUS AX172980 4200 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 107 from Patent WO0142297.
ACCESSION AX172980
VERSION AX172980.1 GI:14597958
KEYWORDS
SOURCE
ORGANISM synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 4200)
AUTHORS Lu, P., Garman, J.D. and Candia, A.F.
TITLE Clasp-3 transmembrane protein
JOURNAL Patent: WO 0142297-A 107 14-JUN-2001;
Arbor Vita Corporation (US)
FEATURES
source
location/Qualifiers
1..4200
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="polynucleotide fragment"
BASE COUNT 1262 a 844 c 907 g 1187 t
ORIGIN
Query Match 60.4%; Score 4123; DB 6; Length 4200;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4136; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
QY 2693 taaatgcttcggaagccttagtaataagcaatccagatatatctggagctccacgctcac 2752
Db 61 TAAATCGTTTCGAAGCGTTAGTAATAGCAATCCAAATATATCTGGAGCTCCACAGCTGAC 120
QY 2753 caagtgtgaagttcgaatcatcatcgcggagtgtaaggtttaagtcgtccaattcctggtg 2812
Db 121 CAGATGATGAGATTCGATCATCATCGGAGTAAGGATTAGATCGCTCCAATTCCTGGG 180
QY 2813 ttaacaatgtgtgtgccaaaagctgcacatggygagatccaaccccaagttcaagtgcagaat 2872
Db 181 TTAACACTGCTGTGTCAAAAGCTGCCCCATGGGAGATCCAACCCCAATCCAAGTCGAGAT 240
QY 2873 caacacagcgtatgagatcgaaagtgttaactcgtatgcttcgcacacagagagctcaagt 2932
Db 241 CAACACAGCGCTATGAGATCGCAAGTTGTAATCGTATGCTTCGACACAGAGACGTCAAGTT 300
QY 2933 tcttcaacaacatcaacggcgcttcaacactaaagaagcttttcaacgagagcttgctt 2992
Db 301 TCTTACAAATTAATACGGGACGCTTACCAACTAAAAAGCTTTTTCACGAGAGCTGCTT 360
QY 2993 tgcagtggtgttcttcagtgycagcgttcgggaatcagcttgcacaagaagctggtct 3052
Db 361 TGCAGTGGGTTGTTTGCAGTGGCAGCGTTCGGGAATCAGCTTTCGCAACAAGCCTGTTCT 420
QY 3053 tttttgaattaaatggttaagaagcatgtgtcacattatcttaacttgaatgaacttgagg 3112
Db 421 TTTTGAATTAATGTAAGAAAGACATGTGCACATTTAATTAATGTAATTAACCTTGAGG 480
QY 3113 ctccaagaaagctgttttccagaacgtttcaatgtagatagcatcttgagatcttgcaaga 3172
Db 481 CTCCAAGAAAATGCTGTTTCCGAACGTTTCAATGATGACATGAGCTCTTGTGACGA 540
QY 3173 cgaatgctagtgatagttcacagattcaagaagagacacagaatggttgagagactca 3232
Db 541 CGATTGCTAGTATATAGTTTACGATTTACAGAAAGACACAGAAATGTTGAGAGACTCA 600
QY 3233 atcaagccttgcatctcttctcaatgacatcgtgtcgtgtttagacagagattgctt 3292
Db 601 ATACAAAGCCTTGCAATCTTTCAAATGATCTGTTGCTGTATGACACAGAGATTGTTT 660
QY 3293 ttgacctataaagctcctctataaacagagtgcttcaagcttactactataacgaatc 3352
Db 661 TTAGCCTTATAAGTCTCTGCTATAACAGAGTGTCTTCAAGCTTTACTATTAACGAATC 720
QY 3353 ccagtgctcgtgtccttgaggcttgatcttctaagaaatcgtcagatcgtacgtacact 3412
```

Db 721 CCAAGTTCCTGGTCTGCTTGAAGCTGGATTTCTACGAATCCTGCAATGATGACACT 780
Oy 3413 atgttaacatlaaacttaacccgcgcagcttaacttaactcaaccgtacatccatccactctg 3472
Db 781 ATGTTACATTAACTTACCTTCGACGCTTACTTACTCCACTCGACTCTCCATCACCTTCTG 840
Oy 3473 ttctcttgcacaacatcctcaagcttctgatttcttaacgaatgtaacaagaccnaaagattg 3532
Db 841 TTTCTTTCGCAACATCTCAGAGTTCTGTGATTTCTACGAATGTACAAAGACCAAAAGATTG 900
Oy 3533 caaataatgttaaatatccgtgccttcgcgaacagcaatatttggcaagacttgct 3552
Db 901 CAAATATGTTGAATTAATCCCTGCTTCGCCCAACAGCATATTTTGGCAGGACTTGTGT 960
Oy 3593 taacagagctgcgtgtaacttttaagacctgactgtaagagacttttggattgataaga 3652
Db 961 TTAACGAGCTGCTGCTCTCTTTTACACCTGTATGCTGAAGGACTCTTTGGATTGATNAGA 1020
Oy 3653 aagtaacataatgtaacacaatttaactctccagtaacagactcaagaccgcgcgtactctg 3712
Db 1021 AAGTCATCAATATGTTACACAAATTTACTCTCCAGTCAAGATCAGACCCGCGTACTCTG 1080
Oy 3713 accctcagataaagctcgtatgagcaatgttcttaactctgactgtaattgattatcatg 3772
Db 1081 ACCCTCAATTAAGGCTCGAGTGGCCATGTGTATCTACTCTGATTGTGTATTCATG 1140
Oy 3773 aaactgtacctcagctgtatatttttaagaactcaacatacgaaggaagacaaattt 3832
Db 1141 AAACGTACCTCAGCTGTATGATTTTACAGAACTCCAAATCAAGAGAAAGAACAAATTT 1200
Oy 3833 gtaataagcagatgatatgaagtgagagcggaagatataaagccaagccgttgcga 3892
Db 1201 GATATGCCACGATGATTTATGAAGTGAGAGCGGAAGTATATAAGCCAGACCGTTGCA 1260
Oy 3893 tggcaatcgcagaggaacatcggtccctcaactcaacaaagcctgcgaattccctcactcgt 3952
Db 1261 TGGCAATGCGAGGGACATCGGTCTCTCAACTAACAGGCTCGGAGTTTCTCTCTCAAGT 1320
Oy 3953 caaagaagtgagcgaacacactactctttcagcagaatcaagtcgaagcctttgactct 4012
Db 1321 CAAACAGTGGCAGGCAACACACTTACCTTTTACGAGATCAAGTCGAGACCTTTTGATCT 1380
Oy 4013 gtaacttttgggttctcaaaaatgcagatgaacagttctcaagaagtgttcaagatc 4072
Db 1381 GTCCTACTTGGGTCTCAAAAATGCAAGATGAAGAGTTCTACAGAGTGTTTACAGATC 1440
Oy 4073 tctcagctcttcagcttaaacccggtactatagatcgtcttactctctgtgtctgtcttg 4132
Db 1441 TCTCAGCTCTGACGCTAAACCGGCTATTAGATCTGCTTATCTCTGCTGCTTGTG 1500
Oy 4133 agtataaagggaaaaaagtgtttgaagaatgaatagcttgaacctttaaagaatacaag 4192
Db 1501 AGTATTAAGGGAAAAAGTGTTTGAACAAATGAATAGCTTGAACCTTTAAGAAATCAAAAG 1560
Oy 4193 acatagagacaaagcttgaagaagctatctctggagcatatgctgcagagcaagaatg 4252
Db 1561 ACATGAGAGCAAAAGCTTGAAGAGCTATTCTTGGAGCATATGAGTCCAGGCAAGAAATGG 1620
Oy 4253 taagcgcaagccgcagagcaagctcgcagaagaagcccatcgcgaagtgcctttggaagtcag 4312
Db 1621 TACGGCGAAGCCGAGAGAGCTCGAGAGAGCCCATCTGGAAGTCCCTTTGGAATCTCAG 1680
Oy 4313 aaaaattgagtgagggaaaagatatgactcaactgctcaaaaacagagaagcttgca 4372
Db 1681 AAAATTTGAGTGGAGAAAAGATATGACTCACTGCGCTCAAAAACACAGAGAGCTTGACA 1740
Oy 4373 aatcaagaagcagagatgaacacgaagacatgattgattgtaaaaccttgcatacagaagca 4432
Db 1741 AATCAAGAGCAGAGATTGAACAGAGCACTGATTGATGAGAAACCTGCTACAGAGCAA 1800
Oy 4433 acctaatcattttaagatacatagagattgtgttccaaacggtttctgttaacggaatcca 4492
Db 1801 ACCTAATCATTTTAAATACATTAGAGATTGTGTTCAAGACCGTTTCTGTAAACGGAATCCA 1860

Oy 4493 aagagagcattcttgg tggagtgctaaagtgctactacacagcatgaccttlaaccaaa 4552
Db 1861 AAGAGAGCAATCTTGGTGGAGTGTCTAAAGTGCTACTACACAGCATGGCTTAACCAAA 1920
Oy 4553 gtcagcttatctacaacactgtttgtctacacagagagccttggtttcaagttctctg 4612
Db 1921 GTGCACTTATCTACAAACCTGTTTGTCTACACAGAGACCTTGGTTTCAAGTTTCCG 1980
Oy 4613 aactctatttgaagaagacagagcaagtgtgattatggcctcagagctctccag 4672
Db 1981 AACTCTTATTTGAAGAGAGACAGACAGTGTCTGATTTATGCTCAGGCTTTCGAC 2040
Oy 4673 acctgagcaatgacatcggtacacatacggtaacacccagtgctccctactactactaa 4732
Db 2041 ACTGTAGCAATACATCGGTACATACGCTCACACCCAGTGCCTTACCTACTATA 2100
Oy 4733 tgaagcaaaacttgaagattgggaataacttgcagaggttlaaaatgcagtaaccaatgt 4792
Db 2101 TGAGGCAAAACTTTGAGATTGGGAATACCTTGGCCAGGGTTAAATGCAGTACCAATGT 2160
Oy 4793 cactatccctcttgggtgggacatctcagaattttaaagaattcttaagagcttctc 4852
Db 2161 CACTATCTCTCTGCTGGGCAATCTCAGAAATTTATAGAAATTTCTTAAGACGTTCTC 2220
Oy 4853 taaagactatattgacataatgctgaagaagatctggaattgagggagaaacaattctctg 4912
Db 2221 TAAAGCAATATGTAGCATATAGCTGAAGAAATCTGGAATTGAGGGAACAAACATTTCCG 2280
Oy 4913 atcagttccagatcgtgatttcaatccacatagtatcttcttga tactgtgaaatga 4972
Db 2281 ATCAGGTCACGAGATCTGTTTCAATCTCCATGATGATTTATGATACAAATTTGCCAAAGTTG 2340
Oy 4973 aggaacccagagagatccctgaatgttgattga tctaattgtacaagaattgcgaaggt 5032
Db 2341 AGGAACACCGAGGAGATCTGAAATGTGATGATATATGATACAAATTTGCCAAAGTTG 2400
Oy 5033 accaagacctctcaga -- -gcga ttgaacctgtgtgcagaacatgcagagcaagcaactcag 5089
Db 2401 ACCAGACCTTCCAGATCTCGATGTGACCTGCTGTGAGAAACATGAGGCAAGCAATCTAG 2460
Oy 5090 aacgaagaatcaatgctgaagctgcagagtgctag tccaactcagacgaactgtgtctg 5149
Db 2461 AAGCAACCAATCATGTGAGAGCTGCACAGTGTCTACTACATCAGCAGCACTGTGTGCTG 2520
Oy 5150 aatatggacatgctggagagaccggaatatctctgtggagatgtgtaaatctcaga 5209
Db 2521 AATATTTGAGCATGTGAGAGACCGGAATATCTCTGTGGATGTGTAACATTTTCA 2580
Oy 5210 atattcatcaatgttttgaagaatctgcgtctcagaatga tctgtatctccagatg 5269
Db 2581 ATATTTATCTATGTTTGAAGAAATCTGCGCTCAAGATGATGATGATGATGATGATGATG 2640
Oy 5270 aagaaglatcgtcctcagaataacttactcgaatcgaatcgaactgtggaattacggaac 5329
Db 2641 AAGAGGTAATCTGCTGGAATAATACTTACTGAGTCAGAGACTTGTGGGATTAATCTGGAAC 2700
Oy 5330 aagaagctgtctctctcctatgctgagca tgfataagaagtaatgaagtttaagaag 5389
Db 2701 AAGCAGTGTCTCTCTCTCTATGCTGTGCGCATGTATGAAGAGATTATGAAGTTTAAAAAG 2760
Oy 5390 tactattcctatcagaaagcaatcogagatgcaagaagaactacacaaatcactcgtga 5449
Db 2761 TACTATTCTCATTTCAATGAGAGCTAATGCGGATGCAAAAGAACTATCCACATTCATGTA 2820
Oy 5450 aactcaagaagcatcgaacaaaatgttca tcaagagttactgctgcgtggagcgaatgttg 5509
Db 2821 AACTCAAGAGCAATTCAGCAAAATGTTTCATCAGATGATGCTGCGGAGCGGATTTTG 2880
Oy 5510 gacactatcttgggtgtgttttataggaaacaaagttcggggagtttggattgaacaaagat 5569
Db 2881 GCACCTATTTTCTGTGTGTTTATGGAACCAAGTTCTGGGGATTTGGATGAACAGAAAT 2940

QY 5570 ttgtttacaaaggagcctgcaataacaaactgtcagagatatctcacagattgagagat 5629
|||||
Db 2941 ttgtttacaaaggagcctgcaataacaaactgtcagagatatctcacagattgagagat 3000
QY 5630 tttaaggagaaagatttggagagagatgtgttgaagtaatcaagaacttaactcttag 5689
|||||
Db 3001 tttaaggagaaagatttggagagagatgtgttgaagtaatcaagaacttaactcttag 3060
QY 5690 acaagtgtaattagatccctaacaaagccatattacagattccatgttggagccact 5749
|||||
Db 3061 acnaagtgtaattagatccctaacaaagccatattacagattccatgttggagccact 3120
QY 5750 ttgacacataatgagatgaaagacagaatccactatttcgacaaaaattacaactctgc 5809
|||||
Db 3121 ttgacacataatgagatgaaagacagaatccactatttcgacaaaaattacaactctgc 3180
QY 5810 gattcatgtactgtacaccccttaacttagatgtgcctgtgccatgaggaaactcatgaac 5869
|||||
Db 3181 gattcatgtactgtacaccccttaacttagatgtgcctgtgccatgaggaaactcatgaac 3240
QY 5870 aattcaaaaggagacccacttctgactacgtctcactgtcttccttatataaagaagg 5929
|||||
Db 3241 aattcaaaaggagacccacttctgactacgtctcactgtcttccttatataaagaagg 3300
QY 5930 tcaatgtcactcataaagaagaagatcacttaacacaaatgaagtgtctataggagaca 5989
|||||
Db 3301 tcaatgtcactcataaagaagaagatcacttaacacaaatgaagtgtctataggagaca 3360
QY 5990 tgcagaaaaaagacacagagagttgtgcattgtcaaacatcaggatccgcagaccccaaaa 6049
|||||
Db 3361 tgcagaaaaaagacacagagagttgtgcattgtcaaacatcaggatccgcagaccccaaaa 3420
QY 6050 tgcctcaagatgtgtactccagagagatctgtagagacacacagttgaataggagcctttgagag 6109
|||||
Db 3421 tgcctcaagatgtgtactccagagagatctgtagagacacacagttgaataggagcctttgagag 3480
QY 6110 ttgcccaagtttctctgtctgaaatcacttagtgaccacaagctcttcagacatcataata 6169
|||||
Db 3481 ttgcccaagtttctctgtctgaaatcacttagtgaccacaagctcttcagacatcataata 3540
QY 6170 aacttgagactctgtctttaaagaatttactaaaaggtgtgtgaagatgaccttaagaanaata 6229
|||||
Db 3541 aacttgagactctgtctttaaagaatttactaaaaggtgtgtgaagatgaccttaagaanaata 3600
QY 6230 aagagcttaattggcgcgggttcaaaaaggagttcaaaaagggaatttggggaactatcttgcg 6289
|||||
Db 3601 aagagcttaattggcgcgggttcaaaaaggagttcaaaaagggaatttggggaactatcttgcg 3660
QY 6290 cttaaaaggagccctacagcccttagatcacagaaagtccctcagttatccaaagccagttat 6349
|||||
Db 3661 cttaaaaggagccctacagcccttagatcacagaaagtccctcagttatccaaagccagttat 3720
QY 6350 gcttgctccctgtgcaagaagattcctcagtcgaatgtgagctttgcgaatgtgactccta 6409
|||||
Db 3721 gcttgctccctgtgcaagaagattcctcagtcgaatgtgagctttgcgaatgtgactccta 3780
QY 6410 aacttaatgtcactgttcttatcatcttgcgaagaagccatgtatccaatgcagatgtgaa 6469
|||||
Db 3781 aacttaatgtcactgttcttatcatcttgcgaagaagccatgtatccaatgcagatgtgaa 3840
QY 6470 aagatctatctggaacacaacatggaatgaaattcctggaatctatctatcatctggaagt 6529
|||||
Db 3841 aagatctatctggaacacaacatggaatgaaattcctggaatctatctatcatctggaagt 3900
QY 6530 cagtggtgcaagaagaataatcaaatgtagatgttgaagcttgagaatcaatgagctatgtctt 6589
|||||
Db 3901 cagtggtgcaagaagaagaataatcaaatgtagatgttgaagcttgagaatcaatgagctatgtctt 3960
QY 6590 ctaagttcgggttaacagagcgttatctttaaagacatttaagacatcaaaagtaact 6649
|||||
Db 3961 ctaagttcgggttaacagagcgttatctttaaagacatttaagacatcaaaagtaact 4020
QY 6650 atacattacatatattatcacatagctlaagttlaaaaatttatcactttaagttcgt 6709

Db 4021 atnatttaccattattttatnaccattacccttaagctttaaatttttttaccctttaaagttcgt 4080
QY 6710 atttcttaattatatacaccattatagattcaatttggagccatttaaatgtagtaat 6769
|||||
Db 4081 atttcttaattatatacaccattatagattcaatttggagccatttaaatgtagtaat 4140
QY 6770 gcttattttaaaggtaataaaatagtgaattgttaccctgtgtgcgcagagcctc 6828
|||||
Db 4141 gcttattttaaaggtaataaaatagtgaattgttaccctgtgtgcgcagagcctc 4199

RESULT 6
AX172978
LOCUS AX172978 3253 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 105 from Patent WO0142297.
ACCESSION AX172978
VERSION AX172978.1 GI:14597957
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
REFERENCE
1 (bases 1 to 3253)
AUTHORS
Lu, P., Garman, J.D. and Candia, A.F.
TITLE
Clasp-3 transmembrane protein
JOURNAL
Patent: WO 0142297-A 105 14-JUN-2001;
Arbor Vita Corporation (US)
FEATURES
source
location/Qualifiers
1..3253
/organism="synthetic construct"
/db_xref="taxon:32530"
/note="polynucleotide fragment"
BASE COUNT 1009 a 633 c 705 g 906 t
ORIGIN

Query Match 47.3%; Score 3230.8; DB 6; Length 3253;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3245; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

QY 3582 aggaactgtgtttaaagagctgtgctgtcatctttagaccctgactgtcgaagactgttgg 3641
|||||
Db 3 AGCACTGTGTGTTTCAAGAGCTGCTGCATTTTAAAGCCCTGATGCTGAAGGACTGTTGG 62
QY 3642 attgcataaagaagtatcaataatgtgtacaacttaactctcagtcacagactcagacc 3701
|||||
Db 63 ATTGCATTAAGAAAGTCAATATGTGTACCAATTTTACTCCAGTCACGACATCACACC 122
QY 3702 ggcgttactgcagccctcagaataaagctcgaagtgcacatgtgttatctacccctgttgg 3761
|||||
Db 123 GCGGTACTGTGACCCCTCAGATTAAGGCTCGAGTGGCCATGTTGTATCTCTGATTGG 182
QY 3762 tatatcatggaactgttaccctcagctgtatgattttaaagaactcacatcaacgaag 3821
|||||
Db 183 TATTATCATGGAAGAAAGTACTCAGCTGATGATTTTACAGAAATCCACATCAACGAGG 242
QY 3822 aagaaccaattgtatgacacactgtatgaatgaagtgagagcgaagtatataagcca 3881
|||||
Db 243 AAGACAAATTTGTATGACCACTGATGATTAAGTGAAGGGAAGTATGATTAAGCCA 302
QY 3882 gacggttgcacatggaactgcgaaggaacatcggtccctcaactaaagaagcctggagttt 3941
|||||
Db 303 GACCGTTGGCATGTGCAATGCGAAGGACATCGTCCCTCAACTACAAAGGCTTGCGATT 362
QY 3942 cctcctcaagtcacaagagtgtgcaggaacacactacatttcaagcagaatcaagtcgaag 4001
|||||
Db 363 CCHCTCAGCTCAACAGAGTGGCAGCAACACTACTCTTTGACGAGATCAAGTCGAAG 422
QY 4002 ccttttgatctgtctactcttgggttctcaaaaatgcagatgaagaagcttcaagaagt 4061
|||||
Db 423 CCTTTGATCTGTCTACTTGGGTTCTCAAAAATGAGATGGAACAGTTCATCAAGAGTG 482
QY 4062 gtttaagatctctcagttcttgagctaaacgagcattagatctgcttactcttctgt 4121

|||||
Db 483 GTTACAGATCTCTAGCTTGCAGCTAAACCGCATATTGATCTGCTTATCTCTGCT 542
OY 4122 gtcttgctttgagatataaaggaaagaaagtgtttgaaacgaatgaatgcttgaccttaa 4181
Db 543 GCTTGTCTTTAGTAAAGGGAAGGAAAGTGTGTAAGCAATGATGCTGACCTTTAA 602
OY 4182 gaatacaaaagacatgagagcaaaagcttgaagaagctatcttcttgagacatggtgcag 4241
Db 603 GAAATCAAAAGACATGAGAGCAAAAGCTTGAAGACCTATTCTTGGAGCATAGGTGCCAG 662
OY 4242 gcaagaaatgtaacgacgaacccgaagacagctcgaagaaagcccatctggaagtgcctt 4301
Db 663 GCAGAAATATGCTACGGCGAAGCCGAGCAGACCTCGAGAGAGCCCACTGCAAGTGCCTT 722
OY 4302 tggaaagtcaagaataattggaagtggagaaagatatgactcaatcggtgcataaacacaga 4361
Db 723 TGGAAAGTCAAGAAATTTGAGGTGAGGAAGATATGACTCACTGGCGCTCAAAACACAGA 782
OY 4362 gaagcttgacaatatcaagagagagattgaaacacgaacacactgattgattgaaacctggc 4421
Db 783 GAAGCTTGACCAATTAAGAGCAGAGATTTGAACAGCACTGATTTGATGGAACCTGGC 842
OY 4422 taacagaagcaaaactcaatcatttagatacatagagaatgtgttgcagaccttctgt 4481
Db 843 TACAGAGCAAAACCTAATCATATTTAGATACATTAGAGATTTGTTGTCAGACCGTTCTGT 902
OY 4482 aacgaatcccaagagagacatcttctgtgtgagtgctaaagaatgctactcaacacagtgcc 4541
Db 903 AACGGAAATCCAAAGAGACATTTCTGTGAGTGTAAATGTCTACTACACACACATGGC 962
OY 4542 ctgtacaagaatgtagtattatctacaacactgtttgtctacacagagagacctgatttc 4601
Db 963 CTGTACCAAGTGAAGTATATCTACACACTGTTTCTACAGAGAGAGCTTGGTTTC 1022
OY 4602 aaagtcttcgaactcttatttgaagaaagacagagcagtgtgctgatttatagtcctcag 4661
Db 1023 AAAGTTCTCTGAACCTTATTTGAAGAGACAGACAGCAGTGTGATTTATGCTCTCAG 1082
OY 4662 gcttctcgaacactgtaacagtagcatgtaacataagtaagtaacaccccaatgtcctcct 4721
Db 1083 GCTTCTCGACACTGTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1142
OY 4722 ttacactaatatgagagcaaaactttagatctggaataacttgcgaagttaaatgca 4781
Db 1143 TTACTTACTAATGAGGCAAACTTTGAGATTTGGGAATTAACCTTGGCAGGTTAAATGCA 1202
OY 4782 ggtaccnaatgtcaactatcctcctctgtgtgggacatctcagaattttaagaatctct 4841
Db 1203 GGTACCAATGTCATATCTCTCTCTGTTGGGCAATCTCAGAAATTTAATGAAGAAATCTT 1262
OY 4842 aagaagctctcraaagacatataatgacatgtaagaaagaaatcaggaatggagaaac 4901
Db 1263 AAGAGGTTCTTAAGAGCATATTTGACATATGCTGAAGAAAGTCTGGAAATGAGGGAAC 1322
OY 4902 aacattctcgtatcagctgcaggaatcgttttcaatctcatalgatcttcttcgtatc 4961
Db 1323 AACATTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1382
OY 4962 tgtgaaatgaagaaacacccagagagatcctgaaatgttgattgataatgataatgataat 5021
Db 1383 TGTGAAATGAAGGAACACGAGAGATCTGAAATGTTGATTGATTCAATGATGATGATGAT 1442
OY 5022 tgcgaaggttaacagagactctccaga---gagattgaacctggtgtgcagaaatgagcag 5078
Db 1443 TGCCAAGGTTACACAGACCTCTCCAGATCTGAGATTGACCTGCTGCTGCAAGATGCGCAG 1502
OY 5079 caagcactcaagaacgaagaacatcgtctgaagctgtaacagtgctcagtcacacagcagc 5138
Db 1503 CAAGCACTCAGAAAGAGCAATCATGCTGAAGCTGCACAGGTGTATGCTCACTCAGCAGC 1562
OY 5139 acttgttctgataatttgagcatgctgaggaacccgaataatctctctgtggagatgtgt 5198
|||||

Db 1563 ACTGTGCTGTAATATTTGAGCATGCTGGAGAGACCGGAATATCTCTCTGTGGATGTGT 1622
OY 5199 aacattcgaataattcactcaatgatttgaagaataatgctggtcgaatgataatgatt 5258
Db 1623 AACATTCAAAATTTCAATCATATGTTTATGAGATCTGGGCTCTGATGATGATGATGAT 1682
OY 5259 atctcagatgaagaagtaatctgctcctgaaatccttacttgatgagtcagagctgtgtg 5318
Db 1683 ATCTCCAGATGAGAAAGATATGCTGCTGGAATAATCTTTACTGAGTCAAGCATGTTGGG 1742
OY 5319 attactggaacaaagcagctgtctcctctctatgctgctgcatgataatgaacgaatga 5378
Db 1743 ATTACTGGAACAAGCAGCTCTCTCTCTATGCTGCTGATGATGATGATGATGATGATGAT 1802
OY 5379 agtttcaaaagtactattctcactatataagtaatctggagtgtaaaagaactctccac 5438
Db 1803 AGTTTCAAAAGTACTATTTCTATTCATGATGATGATGATGATGATGATGATGATGATGAT 1862
OY 5439 aattcattgtaaaacttcaagaagacatctcagcaaatgttcatcagaatgactgctgtgga 5498
Db 1863 AATTCAATGTAATCTCAAGAACCATTTCAAGCAAAATTTGTCATCAGATGATGATGATGAT 1922
OY 5499 gcggatgtttgcaactatlttcgtgtgtgtgttttattgaaaccaagttcggggatttga 5558
Db 1923 GCGGATGTTTGGCACCACTATTTGCTGTGTTTATGGAACCAAGTTGCGGGATTGGA 1982
OY 5559 tgaacaagaattgttttacaagagacatcgaatcaaaactctgcaagaatatactcag 5618
Db 1983 TGAACAAGAAATTTGTTTACAGAGAGCCTGCAATTAACCAATTTGCAAGATATCTCAG 2042
OY 5619 attggaaggaatttctacgagaagaagatttgaagagatgtgtgttgaagtaatacaagact 5678
Db 2043 ATTGAGAGGATTTTACGGAAGAAATTTGGAGAGATGTTGTTGAAGTAAATCAAGACTC 2102
OY 5679 taatcctgtagacaagtgttaattagatctcaacaagacatatatcagattactatgt 5738
Db 2103 TAATCCGTGTGACAGAGTAAATTTAGATCTTAACAAGGATATATTTCAATTAACCTATGT 2162
OY 5739 ggaagcaacttctgacacatatagatgaaagagacgaataacatctatttgcacaaaatta 5798
Db 2163 GGAACCTACTTTGACACATATGATGAAAGGACAAATCACCCTATTTGGACAAAATTA 2222
OY 5799 caatctcgtcgaatcaatgactgtacacaccttactttagatgagcgtgtccatgggga 5858
Db 2223 CAATCTCGTCAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2282
OY 5859 acttcatgacaatcaaaagaaagacatctgactgaatgctcactgcttccctatat 5918
Db 2283 ACTTCATGAACAATTCAAAAGGAAGACCAATTCGACTACGCTCATGCTCTTCTATAT 2342
OY 5919 taaacaagaaggtcaatgtcactcataaagaagagatcatcttaaaccaatgaagtg 5978
Db 2343 TAAACCAAGGCTCAATGTCTCATTAAGAGATCATCTTAACCAATTAAGATTTGC 2402
OY 5979 tattgagacatgcaagaaaaaagacagagagttgagatttgcacaacatcagaatccgc 6038
Db 2403 TATGAGAGCATGAGAAAAGACAGAGAGTTGGCATTGTCACACATCAGAGATCCCG 2462
OY 6039 agaccccaaaatgcttcaagtgttactccagggatctgttagagcaacacagtgtaataagg 6098
Db 2463 AGACCCCAAAATGCTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2522
OY 6099 gacttgaagttgcccaggtttctctgtctgaataactcagtagacccaagaactctcag 6158
Db 2523 GCTTTGGAAGTTGCCAGGTTTTCTGTGTGAATTAACCTAGTACCCAAAGCTCTTCAG 2582
OY 6159 acatcacaataaactggaactctgctttaaagaatttactaaaggltgtgaagaatgctt 6218
Db 2583 ACATCATTAATAACTGGACTCTGCTTTAAAGATTTACTTAAAGGTGTAAGATGATGATGAT 2642
OY 6219 aagaaaaaataagagcttaatttgccggttcaaaagagatatacaaggaatgtggga 6278
Db 2643 AAGAAAAATTAAGGCTTAATTTGGCGGTTCAAAAGGATATCAAAAGGAATTTGGGGA 2702
|||||

QY 6279 actatctgccttaagaagcgcctacagccctacagaaagtccctcagttatcc 6338
 |||||||
 Db 2703 ACTATCTTGGCCTTAAGAAGGCCCTTACACCCCTAGATCCACAAAGTCCCTAGTTATCC 2762
 QY 6339 aagcagatctgctgttccctcgcaagagatccctcagtcgaatgagcttcgcgcaa 6398
 |||||||
 Db 2763 AACCCAGTATGTGCTTCCCTCCGACAGAGATTCCTTACGTGATGAGCTTTCGCAAA 2822
 QY 6399 atgagctcttaactgaatgcactgttttctatctctgcgaagcgcatgtattcaaca 6458
 |||||||
 Db 2823 ATGGATCTCTAAACGTAATGACCTGTCTTTATTCATCTGCAAAAGACCATGTATTCACAA 2882
 QY 6459 tcgagctgtaaaaagatctatcgaaacacacatggaatggaattctcggaattattatcc 6518
 |||||||
 Db 2883 TCGAGGTGTAAGAAAGATCTTTTGAAACCAACATGGAATGGAATTCGTGAAATATTATTATTC 2942
 QY 6519 attgaagaatgcaagtcgcaagaagaataatcaaatgtagatgtttaaagcttgagaatcat 6578
 |||||||
 Db 2943 ATTGAAGATATGCAAGTGGCCAAAGAAATATCAATGATGATGTATTAACGCTTGAGATCAT 3002
 QY 6579 ggctatggcttctaatgttcggttaacaaagctgtatctctttaaagacattttaaagctc 6638
 |||||||
 Db 3003 GGCTATGTTCTTAATGTTGGGTAAACAAGCTGTATCTTTTAAACAATTTTAAATGACTC 3062
 QY 6639 aaagctacactatacattaccattattatacacaagtaagtaagtaaaaattattacac 6698
 |||||||
 Db 3063 AAAGTACACTATACATTATACCATTTATATCCATAGCTAAGCTAATAAATTTATTTCAC 3122
 QY 6699 cttaagctcgatcttcttaattatacaaccattatagatcattcttgagaccatttta 6758
 |||||||
 Db 3123 TTTAATGTTGTAATTTTAAATTTATATACCATTTATAGATTCATTTTGAGACCATTTTAA 3182
 QY 6759 aatgtgatactcttattttaaaggtactcaaaaataatgtgaatgttactcgtgcgcg 6818
 |||||||
 Db 3183 AATGATGATATCTTATTTTAAAGGTAATAAATAATGATGATGATTTTAACTCGTGGCGC 3242
 QY 6819 ccagagcctc 6828
 |||||||
 Db 3243 CCAGGGCCCTC 3252
 RESULT 7
 AK055905 3327 bp mRNA linear PRI 31-OCT-2001
 LOCUS Homo sapiens CDNA FLJ13143 fis, clone MESAN1000101, weakly similar
 DEFINITION to Rat. try gene product.
 ACCESSION AK055905
 AK055905.1 GI:16550749
 VERSION oligo capping; fis (full insert sequence).
 KEYWORDS Homo sapiens normal mesangial cells (NHMC56046-2) cDNA to mRNA,
 SOURCE clone.lib:MESANI clone:MESANI000101.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Tashtiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
 Fujizumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,K.,
 Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
 Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
 Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
 Watsutsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
 Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
 Nagahari,K., Masuko,Y., Nagai,K. and Isogai,T.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 3327)
 AUTHORS Isogai,T., Otsuki,T. and Sugiyama,T.
 TITLE Direct Submission
 JOURNAL Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
 Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 252-0612, Japan
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan: cDNA full insert sequencing:
 Research Association for Biotechnology (RAB): cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 key technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: RAB and
 HRI.

FEATURES source

1..3327
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MESAN1000101"
 /cell_type="normal mesangial cells (NHMC56046-2)"
 /clone.lib="MESANI"
 /note="cloning vector: pME18SFL3-primary culture, normal
 mesangial cells"
 54..2279
 /note="unnamed protein product"
 /codon.start=1
 /protein.id="BAB71042.1"
 /db_xref="GI:16550750"
 /translation="MFELSVPRQOHYLAGLVLTDELAVLPDPAEGLEGHKKVIMNV
 HNLSSHSDPRYSDPQIKARVAMLYPLIGIMETVPOLYDFETTHNGRPICIAT
 DYESESGSMISQTVAMAIAGTSVPQTRPGSEFLTSTGROHTTFSAESSRLICL
 LMYLKNADEVTLQKMTFDSLVLQNLRLDLTLVLCVSCFEYKKGKVFERNLSLTFKSK
 LDKSRAEIEHEALIDGNLTLEANLILDTLELVQTVSTESKESILIGVGLVLLSHM
 ACNQSVAVYLQHCFAIQRALVSKFPELLFEETEQCADLCLRLLRCCSSIGTIRSHAS
 ASLYLLMRONFELGNFARVKKQVYKMSLSLQVDFEEMFNEFLRSLRTILYAEEDL
 ELSETTFPOVOVLVNLHMLISLVYKMEHODEPMLIDRYLRAKGVQTSPLRLT
 WLONMAGKHSERNHAEAOCLVYSALVAEYLSMLEDKRYLPVCVFQNTSPDLRLT
 ESNVSDVVSPEDEICSGKVFESGLVALLFOAAASPMACMYAVNVEYVLLPIH
 EANRDAKKISTIHGKIQEAPFKSTIVHODKRMHGCTTFRFGVFTKRGDDLEDFVKEP
 ATKLEAIEISHRLEGFYGERFGEDVEVIVINDSPVDCIKLDPMTHTIFRLPMWSHTLTHM
 R"

CDS

BASE COUNT 1046 a 640 c 711 g 930 t
 ORIGIN

Query Match 46.9%; Score 3201.8; DB 9; Length 3327;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 3285; Conservative 0; Mismatches 32; Indels 14; Gaps 5;

QY 3485 catctcagagcttcgagcttctcagaaatgtaagaacaaagatgtgaatatgttg 3544
 |||||||
 Db 1 CATCTCAGAGTCTGGGATTTTCTACGATGTACAAAGCCAAAGATTGCAATATGTGTG 60
 QY 3545 aattatcgtgcctcttcgcgcaaacagcatattctgacagactgtgttaacagagctgg 3604
 |||||||
 Db 61 AATPATCCGTCGCTTCCGCAACAGCATATTATTTGCGAGACCTGTGTTAAACGACTCG 120
 QY 3605 ctgtcattttaagaccctgatacgtcgaagactgtttgagatgcatagaagaatcatcaata 3664
 |||||||
 Db 121 CTGTCAATTTTAAAGCCCTGATGCTGAAGGACTGTTGTGATGTGATTAAGAAAGTCATCAATA 180
 QY 3665 tggtaacaattactcttcagtcacagactcaaaccccggtactctgccccagagata 3724
 |||||||
 Db 181 TGGTACACAATTTACTCTCCAGTCACAGCTCAACCCGGCGTACTGTGACCCTCAATATTA 240
 QY 3725 aggcctcagtcgacatgttatacactcctgtatcttgatatacagagaaactgtaccctc 3784
 |||||||
 Db 241 AGGCTCAGTGGCCATGTTGTACTACTCTGTGATTTGATATCAATGAAACGTATCCTC 300
 QY 3785 agctgtatgatttaacagaactcaacaatcaacagagagaaccattgtatagccactg 3844
 |||||||
 Db 301 AGCTGTATGATTTTACAGAAACTCACAAATCAACAGAGAACCAATTTGTATAGCCACTG 360
 QY 3845 atgattatgaagtggagggagagatgataagcgaagcgtttgcataggaatctgcgag 3904
 |||||||
 Db 361 ATGATTTATGAAGCTGAGGCGGAAGTATGATTAAGCCAGACCGCTTGCCATGCGCAATGCGAG 420
 QY 3905 ggaacatcgatccctcaactacaagccttcgagcttccctccctcaacgatacagagtgaga 3964
 |||||||

Db 421 GGACATCGTCCCTCACTAACAGAGCCCTGGCAGTTTCCTCCTCACTCAACAGATGGCA 480
OY 3965 ggaacacactaactcttcagcagaatacaagtgcgaagcccttttgatctgctacttgagg 4024
Db 481 GGACACACCTACCTTTTACGACAGATCAAGTCGAAGCCCTTTGATCTGTCTACTTTTGGG 540
OY 4025 ttctcaaaaaatgcagatgaagaagctctacagaagtggtttacagatctctcagctctgc 4084
Db 541 TTCTCAAAAATGCTCAGATGAACAGTGTCTACAGAAGTGTTTACAGATCTCTCAGCTTGGC 600
OY 4085 agctaaacccggtatagctctgctcttactctctgtgtctgtctcttgatgaataaggga 4144
Db 601 AGCTAAACCGGCTATTGATGATGCTTTATCTCTGTGTGTCTTGGCTTTGATGATTAAGGGA 660
OY 4145 aaaaaatgcttgaacgaagaataagcttgacctttaagaatatcaaaagaacatgaagaaga 4204
Db 661 AAAAATGTTTGACCAAGATGATAGCTTGACCTTTTAAGAAATCAAAAGACATGAGAGCA 720
OY 4205 agcttgaagaagctatctcttgaggacataagtgccaggccaagaatggtaacggaagcc 4264
Db 721 AGCTTGAAAGAACTATTCTTGGGACATAGTGCCAGGCAAGAAATGGTACGGCGAAGCC 780
OY 4265 gaggagaagctcgaagaagaagccatctggaagatgctccttggaagtgtaagaataattgaggt 4324
Db 781 GAGGACAGCTCGAGAGAACCCCATCTGGAAAGTGCCCTTGGAAAGTCAAGAAATTTGAGGT 840
OY 4325 ggaagaaagatatgactcactgctgcacaaacacagaagaagcttgacaataagaagcag 4384
Db 841 GGGGAAAGATATGACTCTGCTGCTCAAAACACAGAGAGCTTACAAATCAAGAGAG 900
OY 4385 agattgaacagaagaagcatttgatggaagacgttgctacagaagaacaaactaatcattt 4444
Db 901 AAGATTGAACACAGACACAGATTTGATGGAACCTGGCTACAGAAACCAACCTAATCATTTT 960
OY 4445 tagatacatatagaattgttgcagaagccgttctgcgaagaatccaaagaagcattc 4504
Db 961 TGATATCATTTAAGATTTGTTGTTGACACCGTTCGTAAACGGAATCCAAAGGACCATTC 1020
OY 4505 ttggttgagatgctaaagatgctactacacagaagcattgacccaagtgcaagttatc 4564
Db 1021 TTGGTGAAGTGTAAAGTGTACTACACAGATGGCCTGTAAACCAAGTGAAGTTATTC 1080
OY 4565 taacaacactggttgcatacagaagaagccttggttcaaaagttccctgaactctatttg 4624
Db 1081 TACACACACTGTTTCTTACACAGAGAGCCCTTGTTCAAAAGTTTCTCAACCTTATTTTG 1140
OY 4625 aagaagagaacagagcagtgctgtaattatagctcagaagctcttcgaagcgttgaacata 4684
Db 1141 MAGAAGACAGACAGAGTGTCTGATTTATGCTTCAAGGCTTCTCGACACTGTAGCAGTA 1200
OY 4685 gcatcgatacaatacgttccacacccacagtgcccttactactactaatatgagcgaact 4744
Db 1201 GCATCGGTTCATATACGTCACACGCCAGTGCCTCCCTTACTACTATATGAGGCCAAACT 1260
OY 4745 ttgaaattgggataactcttcgaagtgtaaaatgcaggtaccaaattgactatctcct 4804
Db 1261 TTGAATATGGGAATTAACCTTGGCAGGTTAAATCAGAGTAACATGTCACTATCTCTCT 1320
OY 4805 tggctggcacatcacaagaatttaagaagaattcttaagcgttcttaagaagctat 4864
Db 1321 TGGTGGGACATCTCAGAAATTTTATGAAGATTTCTTAAAGAGTCTTAAAGACTATAT 1380
OY 4865 tgacataatgcgaagaagatctggaattgagggaacaacattctcagtcagtcagag 4924
Db 1381 TGACATATGCTGAAGAAGATCTGGAATTGAGGAAACAACATTTCTGATCAGGTCCAGG 1440
OY 4925 atctggtttcaatctccatagatctcttcttgataactgtgaaatgaagaacaacagg 4984
Db 1441 ATCTGTTTTCATCTCATATGATTTCTTCTGATACGTGAAATGGAAGAACCCAGG 1500
OY 4985 aggatcctgaagaatgcttgatgtaactgaatgaagaattgccaaggttaccagaagcttc 5044
Db 1501 AGGATCTCGAAATGTGTGATGTATTAATGTACGAATTTGCCAAGGGTTACCAAGACCTCTC 1560

OY 5045 caga---gcgaatgacctggttgacagaaatgagcaggcaagcaactgaagaacgaacatc 5101
Db 1561 CAGATCTGCGACTACCTGGTTGCGAACAATGGAGGCAAGCACTCGAAGCAATC 1620
OY 5102 atgctgaagctgcacagtgctctagtcacacacgaagcacttgctgctgaatcttgagca 5161
Db 1621 ATGCTGAAGCTGCACAGTGTCTATCCACAGAGACACTTGTGCTGATATTAGCA 1680
OY 5162 tgctggaagcaggaatatctcctgttggaatgctgaacatctcagaatatttcaacta 5221
Db 1681 TGCTGGAGGACCGGAATATCTTCTGTGGATGTGTAACTTTGAAATTTCAATCTA 1740
OY 5222 atgcttgaagaatctgcggtctcagaatgtagtcatctccaagtgaagaaggtatct 5281
Db 1741 ATGTTTGAAGAAATCTGCGGTCTCAGATCATGATGATGATCCAGATGAAGAAGTATCT 1800
OY 5282 gctctggaataatcttactcagtcaggaagcttgaggatctacggaacaagcagctgctt 5341
Db 1801 GCTCTGGAATAACTTTACTGAGTCAGAGCTGTGGGATTTACTGGAACAAGCAGCTGCTT 1860
OY 5342 ccttctctatgctgcacatgtaagaagcatttaagaatttcaaaagtaacttactcta 5401
Db 1861 CCTTCTATGCGCTGGCATGTATGAACAGTAAATGAAGTTTCAAAAGTACTTATTCCTA 1920
OY 5402 ttcaagaagcctaactcggagatgcgaagaacatccacaattctgtraaacttcaagaag 5461
Db 1921 TTCTGTAAGCTAATCGGAGATGCCAAGAACTATCCAAATTCATGGTAACCTTCAAGAAG 1980
OY 5462 catcagaacaaatgttcatcagaagtaactgcttggaagcagatgcttgaccattctc 5521
Db 1981 CATTCACCAAAATGTTTCATCAG-----GATGGTAAAGCGGATGTTGGCACCATTATTTTC 2034
OY 5522 gtgtgtgttttataggacaagaattcggggaatttgatgaacaagaattgcttacaagg 5581
Db 2035 GTGTTGTTTTTATGGAACCAAGTTCGGGATTTGGATGAACAAGATTTGTTTAAAGG 2094
OY 5582 agcctgcaataacccaacttgcagagatactcacaagaattggagggaattttagaggaa 5641
Db 2095 AGCTGCAATATACCAAACTTGCGAGATATCTCAGAGATTGGAGGATTTACGGGAAA 2154
OY 5642 gatttgaagagatggtgttgaagtaacaaagacttaactcctgtlagacaagtgtaat 5701
Db 2155 GATTGGAGAGGAGTGTGTTGAAGTATCAAAAGCTATATCTGTATGACAAAGTAAAT 2214
OY 5702 tagatcctaagaagcatalattacagattacatctatgtggagccaacttgaacatatg 5761
Db 2215 TAGATCTTAAC-ACGCATATATTCAGATTAACATGTGGAGCCATACCTTTGACACATATG 2273
OY 5762 agatgaaggaacagatcaactatttcgacaaaatatacaactctcgttcgattact 5821
Db 2274 AGATGAAGGACAGATACACTATTTCGACAAAATTAATCTTCGTCGATTTATGTACT 2333
OY 5822 gtacaccttactttagatggtgcgtgcccagtggaacttcaagaatctcaaaaagg 5881
Db 2334 GTACACCTTTACTTTTGAATGGCCGTGCCATGGGAACTTCATGAACAATTCAAAAGGA 2933
OY 5882 agaccattctgactaagcttcacatgctcttccctatatataaaagaaggtcaatgtaactc 5941
Db 2394 AGACCAATTCGACTACGTCTCATGCTCTTCTTATATTAAGAAAGGTCAAAGTACTC 2453
OY 5942 ataaagaagatcatcttcaacccaattgaagtctctatcttgaggacatgcaagaataa 6001
Db 2454 ATTAAGAAGAGATCATCTTAAACACCAATTGAATTCCTATTGAGGACATGCAAAAAA 2513
OY 6002 cacaggaattggcaattgcaacaacatcagaatcccgcaagaccccaaaatgcttcagatg 6061
Db 2514 CACAGGAGTTGGCATTTGCAACACATCAGGATCCCGACAGCCCAAAATGCTTCAATGG 2573
OY 6062 tactcagagatctgttaggacacagtgtaataagggtccttggaagtgtgccaggtt 6121
Db 2574 TACTCCAGGATCTGTAGGCCACCAAGTGAATCAGGGGCTTTGGAAGTTGCCACAGGTTT 2633

QY 6122 ttctgtcgaatactactagtagaccgaagctcttcaagacatcataataaactgacctct 6181
|||||
Db 2634 ttctgtcgaataatctactgtagaccgaagctcttcaagacatcataataaactgacctct 2633
QY 6182 gctttaaagatttctaaagagtgtagaagtcctttaaagaanaataaagcttaattg 6241
|||||
Db 2694 gctttaaagatttctaaagagtgtagaagtcctttaaagaanaataaagcttaattg 2753
QY 6242 ggcgggttcaaaaaggatgatacaaaaggaaattg--ggaaactatcttcgccttaagaagg 6299
|||||
Db 2754 ggcgggttcaaaaaggatgatacaaaaggaaattg--ggaaactatcttcgccttaagaagg 2813
QY 6300 cccctaaagccctagatc--caagaagcctcctagatgatacaagccagatcttgcttgcc 6357
|||||
Db 2814 cccctaaagccctagatc--caagaagcctcctagatgatacaagccagatcttgcttgcc 2873
QY 6358 cctgcacagagatcctcctcagtcgaatgagcttcgcgaataatgatactcaaatgaat 6417
|||||
Db 2874 cctgcacagagatcctcctcagtcgaatgagcttcgcgaataatgatactcaaatgaat 2933
QY 6418 gcaattgtttattcatctgcgaagagccatgattatcaacatcgagtgtagaagaatccta 6477
|||||
Db 2934 gcaattgtttattcatctgcgaagagccatgattatcaacatcgagtgtagaagaatccta 2993
QY 6478 ttggaaaccaaatggaatggaattctggaataattatattcatggaagaatcgagtgacc 6537
|||||
Db 2994 ttggaaaccaaatggaatggaattctggaataattatattcatggaagaatcgagtgacc 3033
QY 6538 aagaataatacaaatgtagatgtagtaacgcttggaataatcagatgtagtattcctaattg 6597
|||||
Db 3054 aagaataatacaaatgtagatgtagtaacgcttggaataatcagatgtagtattcctaattg 3113
QY 6598 cgggtgaacaagctgtatctttaaagaacatttaataagctcaaaagtgatacactatacat 6657
|||||
Db 3114 cgggtgaacaagctgtatctttaaagaacatttaataagctcaaaagtgatacactatacat 3173
QY 6658 accattattatcacatagctaaagttaaaattatcattcacttaagtcgattatttcta 6717
|||||
Db 3174 accattattatcacatagctaaagttaaaattatcattcacttaagtcgattatttcta 3233
QY 6718 attatatacattatagattatcatttggaaacattttaaattgtagtaattgcttatt 6777
|||||
Db 3234 attatatacattatagattatcatttggaaacattttaaattgtagtaattgcttatt 3293
QY 6778 taaagttactaaaataatgtagaattgcttac 6808
|||||
Db 3294 taaagttactaaaataatgtagaattgcttac 3324
RESULT 8
AK05493 2452 bp mRNA linear PRI 31-OCT-2001
LOCUS Homo sapiens cDNA FLJ30931 f1s, clone FEBRA2006873, weakly similar
DEFINITION to Trg protein.
ACCESSION AK05493
VERSION AK05493.1 GI:16550228
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens fetus brain cDNA to mRNA, clone_lib:FEBRA2
Clone:FEBRA2006873.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (sites)
AUTHORS Nishi,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
Kikuchi,H., Kanda,K., Wagatsuna,M., Murakawa,K., Kanehori,K.,
Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
Sugano,S., Nagahari,K., Masuno,Y., Nagai,K. and Isogai,T.
TITLE NEBO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2452)

AUTHORS Isogai,T., Otsuki,T. and Sugiyama,T.
TITLE Direct Submission
JOURNAL Submitted (24-Oct-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
COMMENT NEBO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan: cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: RAB and
HRI.
FEATURES
Source Location/Qualifiers
1..2452
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="FEBRA2006873"
/tissue_type="brain"
/clone_lib="FEBRA2"
/dev_stage="fetus"
/note="Cloning vector: pME18SFJ3"
171..2051
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB70933.1"
/db_xref="GI:16550229".
/translation="MACNONAVYLQHCFAOTRALVSKPELLFEETFOCADLCLRL
RSCSSIGTRISHASASLYLKRQNEFEIGNPFAVKMOVWMSLVSLSVTFONFNEFL
RSLKTIILYAEDELEIRETTPPDQVDLVFRLHMLISDTYVKMHEQSDPMLIDLMY
RLAKGQTPSDRLTFLWLOMACKHSERSHSAALVSAALVLALEYLSMLEDKRYLP
VGCVFQNTSNVLESASVDVVPDEEGICSGKYFTSGSLVGLVLEQASFSNAGM
YKAVNPNYKVLPIHEANRDAKSLSTJNCKLOEASKIVHSTGEMRPFNGYFRGKY
GTRFDLDEQKTVKEPATLTKLAELSHRLEBGTGREGREDDVYEVKSDNPDKCLDP
NKAVYIQITVEYPTVEKDKITVFDKNYNLRKMYCTPTFLDSRAGEVLEHQRK
TLITTSHPYIKTRVNVNKHKEIILTPLEVAIEDMKRQLEAFHQDPAQDMQLQ
NYLQSGVGTAVNGPLEVAOVLEISPPSKLFRHNKRLRCFKDFTKRCBDALRNK
SLIGPQKREYQRELEERNYHRLKEALQPLINKRIPQLYKAVLPVTCRDSFSRMSLRKM
DL"
BASE COUNT 781 a 458 c 509 g 704 t
ORIGIN
Query Match 34.6%; Score 2360.2; DB 9; Length 2452;
Best Local Similarity 98.6%; Pred No. 0;
Matches 2414; Conservative 0; Mismatches 28; Indels 7; Gaps 3;
QY 4367 ttgacaaatcaagaagcagagattgaaacgaagcactgattgtagaaccctggtctacag 4426
|||||
Db 1 TTGACAAATCAAGAGCAGAGATTGAAACGACGACACTGATTGAGAAACCTGGCTACAG 60
QY 4427 aagcaaaccttaatcatttagatagatgagattggtttcagaccgtttctgtaagc 4486
|||||
Db 61 AAGCAAACTATCATTTTGTGATACATTAGAGATTTGTGTTCAGACCTTTCTGTACGG 120
QY 4487 aatccaaagagagacattctggtggaagtgcctaaagtgcctacacgaacaatgagctgta 4546
|||||
Db 121 AATCCAAAGAGACATTCTTGGTGAGCTCTAAAGTGTCTACACAGCATGCGCTTGA 180
QY 4547 accaaagtgcagttatctacaacactglttgcctacagagagccttggtttcaagt 4606
|||||
Db 181 ACCAAATGCAAGTTATCTACAAACACTGTTTCTACACAGAGAGCCTTGGTTCAAAAT 240
QY 4607 ttctgtaaccttatattggaagaagagacgaagagtgctgtagttatgacctcagcttc 4666
|||||
Db 241 TTCTGTGAACCTTATTTGAAGAAGAGACAGAGAGTGCTGTATTTATGCTCAGGCTTC 300
QY 4667 tccgacactgtagcagtagatcggtatacaatcagtcacaccccaagtgctcccttacc 4726
|||||
Db 301 TCCGACACTGTAGCAGTAGCATCGGTACATACGGTCCACAGCGCAATGCTCCCTTACC 360
QY 4727 tactaatgaagcaaaacttggaattggaataacttgcgaaggtttaaattgaggtac 4786
|||||

Dh 361 TACTAATGAGCAAAACTTTGAGATTGGGAATTAACCTTGGCAGGGTTTAAATGCAAGTTAA 420
Qy 4787 caatgcatacctcctctgtgtggcacatccagaaatttaataagaatcttaagac 4846
Dh 421 CAATGTCACTATCCCTCGTGGGSCACATCTCAAAATTTTAATGAAGAAATTTCTTAAGAC 480
Qy 4847 gtctcctaagaatactatgacataatgctgaagaatctggaattgaagggaacaacat 4906
Dh 481 GTTCTCTAAAGACTATATTTGACATATGCTGAAGAAGATCTGGAATTTGAGGGAACAACAT 540
Qy 4907 ttcctgacatcaggtccagatcgtgtttcaatctccatataatcttctctgatactgtga 4966
Dh 541 TTCTCGATCAGGTCCAGATCTGGTTTCAATCTCCATATATATCTTCTATCTGGA 600
Qy 4967 aaatgaagaacacacaggagagatcctgaatatgtgattgatacttaatgtaacgaattgca 5026
Dh 601 AAATGAAGGAACACAGAGAGATCTCTGAATTTGATTTGATTTGATTTGATTTGCA 660
Qy 5027 aggtttaccagaactcaccaga--gagattgacctgtgtgacagaacatggaagcaagc 5083
Dh 661 AGGGTTACCAGACTCTCAGATCTGCGATTTGACCTGGTTGCAAGAACATGSCAGGCAAGC 720
Qy 5084 actcagaagcagaacataatgctgaagctgacagtgctgatacctcaccagcaacttg 5143
Dh 721 ACTCAAGAGGAAGCAATATGCTGAAGCTGCAACAGTGTCTAGTCCACACACACACTTG 780
Qy 5144 ttgcgtaatatttgacatgctgtgagacccgaaatacttcctctgtgagatgttaacat 5203
Dh 781 TTGCGAATATTTTGAGCATGTCTGGAGGACCGGAATATCTTCCCTGTGGGATGTGAACAT 840
Qy 5204 ttcaagaatacttcaataatgttttaagaagaactcgggtctcagatgaatgtgatalctc 5263
Dh 841 TTCACAATATTTCACTATATGTTTGAAGAAATCTGCGTCTCAGATGATGTGATCTC 900
Qy 5264 cagaagaagaagatctcgcctctgnaaaatacttactacgaatcgaactgttgatctaac 5323
Dh 901 CAGATGAAGAGAGATATCTGCTCTGAAAAATCTTACTGAGTCAAGACTTGTGGATTTAC 960
Qy 5324 tggaaacaagcagctgctcctcctcctatgctgtgacatgaagaacagtaataagttt 5383
Dh 961 TGGAAACAAGAGCTCTCTCTCTATGCGTGGCATGTATGAAGCACTTAATGAAGTTT 1020
Qy 5384 acaaatccttaattcctatcctaagaactaaatcggaatgcaagaacaaactcacaacttc 5443
Dh 1021 ACAAAGTACTTATCTTCTATGAGAACTAATCGGATGCAAGAAACTATCCAAATTC 1080
Qy 5444 atggaatactcaagaacatctcagcaaatgtttcatcaagaatctggtcggagcggga 5503
Dh 1081 ATGGTAACTTCAAGAACATTCACGAAATTTGTCATCAGAGTAAGTGGGAGCGGA 1140
Qy 5504 tgtttggaacctatttcgtgtgtgttttttaagaaccaagtctcggaatttgatgaac 5563
Dh 1141 TGTTGGCACCGCATTTTGTGTGTGTTTATGGAACCAAGTTGGGGATTTGGATGAAC 1200
Qy 5564 aagaatttgtttaagaagagccgtcaataacaaactgcaagatatctcagagatttg 5623
Dh 1201 AAAAATTTGTTTACAGAGAGCTCGCAATTAACCAAACTTGACAGAGATATCTCAGATTTG 1260
Qy 5624 agggatttttaaggagaagaatttggaagagatgtgtgtgaagtaatacaaaccttaac 5683
Dh 1261 AGGGATTTTACGGAAGAAAGATTTGGAGAGATGTGGTGAAGTAATCAAAACCTTAATC 1320
Qy 5684 ctgtagacaagtgttaaatltagatccctaaagaagcatatattcagattaccatgtgagc 5743
Dh 1321 CTGTAGACACAGTAAATTAATAGTCTTAACAGGCATATATTCAAGATTATGTGAGC 1380
Qy 5744 catacttgaacacataatgaatgaagaagacaatcaactatttcgacaaaataataaac 5803
Dh 1381 CATACTTTGACACATATGAGATGAAGGACAGAAATCACTATTTTGACAAAATTAACAATC 1440
Qy 5804 ttgcgtcgaatctgactgtaacaccttactttagatggcgcgtggccatgagggaacttc 5863
Dh 1441 TTGCTGATCTCATGTACTGTACACCCCTTACTTATGATGGCGGTGCCATGGGAACCTTC 1500

Qy 5864 atgaacattcaaaagaagaccatctcactacgtctcactgcttcccttaataaa 5923
Dh 1501 ATGACATATTCAAAAGAGAGACCATTTGACTGACTGACTGCTCTGCTTCTTATATTTAAA 1560
Qy 5924 caaggtlcaatgtcactcactaaagaagatcacttaacacccaattgaagtgtatgt 5983
Dh 1561 CAAGGTCATATGCTCATTAAGAGAGATCATTTTAACCAATTTGAAGTGTGATTTG 1620
Qy 5984 aggaacatgcagaagaagacacaggaattggtcatcttgcaacatcagagatcccgagac 6043
Dh 1621 AGGACATGACGAAAAAGACACAGAGATTGGCATTTTGCAACATCAAGATCCCGAGACC 1680
Qy 6044 ccaaaatgctcagatcgtactcagaagatcgttaagacacagatgaatcgaaggacct 6103
Dh 1681 CCAAAATGCTTACATGATGTAATCTCAGGATCTTAAAGCACACAGATCAAGGACCTT 1740
Qy 6104 tggaaattgcccaggttctcgtctgaatactagtagaccacaaagctctcagacatc 6163
Dh 1741 TGGAAAGTGGCCAGGTTTCTGTGAAATACCTAGTGAACCAAGCTCTTACAGATC 1800
Qy 6164 atataaactgagcactcgtctttaaagatttacttaaaagtgtagaagtgccttaaga 6223
Dh 1801 ATATATAACTGCACTCTGCTTAAAGATTTTACTTAAAGGTGTGAAGATGCTTAAAGAA 1860
Qy 6224 aaaaagaagcttaattggtgcggtlcaaaagaagatcacaaggaatg--gggaact 6281
Dh 1861 AAAATTAAGACTTAATTTGGCCCGGATCAAAAGAGATCAAAAGGAACCTGGAGAGAACT 1920
Qy 6282 acttcgctttaaagagccctcacaagccctagatca--cagaagtcctcagttatcca 6339
Dh 1921 ATCATCCCTTTAAAGAGGCCCTACAGCCACTGATCAACAGAAAGATCCCTCAGTTATCA 1980
Qy 6340 agccagatgctgtgtccctgtccagaagatctccttcagtggaatgaaacttcgnaaa 6399
Dh 1981 AGGAGTATTTGCTGTCACTGTCACAGAGAAATCTTCACTGCAAGAACCTTCCGAAAA 2040
Qy 6400 tggatcctaaactgaatgtaactgtttatcatctgcaaaagacagatcttaacaat 6459
Dh 2041 TGGATCTCTAACTGAATGATGCACTTGTTTATTTATCTGCAAAAGGCCATGTATCAACT 2100
Qy 6460 cgaagtgaagaagatcctatggaacccaatggaatggaatcttgaataattatcca 6519
Dh 2101 CGAGTGTGAAGAAATGATCATTTGGAAGAACAGCATGGAATGGAATTTATTTATCA 2160
Qy 6520 ttgaagaatgcatgtgcccagaagaataataatgtagatgttgaagccttggaatcatg 6579
Dh 2161 TTGAAGATGCAAGTGGCAAGAAATATCAAAATGTGATTTGTTAAAGCTTGAAGATCATG 2220
Qy 6580 gctatggttctaatgttgcggtcaagaagctgtatcctttaaagacatttaagactca 6639
Dh 2221 GCTATGTTCTATATGTCTGTGTAACAGCTGTATATCTTTTAAGACATTTTAATGACTCA 2280
Qy 6640 aaggtacacatacatltaacattatltacacatagcctaaggttlaaaaaattatctact 6699
Dh 2281 AAGTACACTATACATTTTACATTTATTTATACATGCTAGTGAAGTAAATTTATTTCACT 2340
Qy 6700 ttaagtcgattttttaaattatataccacatttaagattcaatttgaagccattttaa 6759
Dh 2341 TTAAGTTCGATTTTAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 2400
Qy 6760 atgtagaatcgtctattttaaaggtactaaagaataatggaagtttac 6808
Dh 2401 ATGTAGTAATGCTTATTTTAAAGGTACTATTAATATGTGAATGTTTAC 2449

RESULT 9
AX173175 AX173175 6372 bp DNA linear PAT 03-JUL-2001
LOCUS DEFINITION Sequence 1 from Patent WO0142295.
ACCESSION AX173175
VERSION AX173175.1 GI:14598036
KEYWORDS

OY 1286 tagaatacagtaactcggagaaacgaaaggtctctgtgtaagagaggaatcttagtatgtg 1345
DB 1285 AG-----GGCAGAGCGCGGCCAGCCTGGACAGAGCCGCG----- 1298
OY 1346 ttggcagacgatactcaattbaaagacaacaagttgagatgtagctgttaactttagcagct 1405
DB 1299 -CCGTGGGGGGCCCCAGACCGGGGCGATAGTGGGAGAGCGCTGGACGCTTCTTGACT 1357
OY 1406 ttgcacagactactctccacagtgacaatatttttaagcaagaaagagacccgcttaagt 1465
DB 1358 TCCGTCACAGCCACGCTACTGTCAACAACCTTTAAAGCAGAGAGCTGAGAGACTCAAGTG 1417
OY 1466 atgaagatctctacaatctctgtgatatgagaagggcaactctgtcttaccgagac 1525
DB 1418 ACGAGAGCACTCTTCAGATTCTGCTGACATGAGGCGCCGCTGCTCCCTGCTGGCGGAC 1477
OY 1526 taagaccattatacagctcagctcaagataagatacttctccgcacccgaaatcccatc 1585
DB 1478 TACGTCTGTACTGCTCCAGCTCAAGATCGACATTTCGCGCTCTGAAATCCCACT 1537
OY 1586 attgcatactccgagctgtctcaagtgaaagcttaccgagagtagagttagactta 1645
DB 1538 TCTGCTCTCCCTGAGCTGCTGCTCATATCAAGCCCTACCCGAGCCCAAGGGGCGGCCCA 1597
OY 1646 ccaagaaatctttagagttcccgcaaggaagtattatgttccaaagactaattacagaa 1705
DB 1598 CCAAGGAGATTCTGGAGTTCCCGCCGCGAAGTCTATGCCCCCATACCAAGTACAGGA 1657
OY 1706 attctctataataaccctcaagagcttaatttgcacaatcgtccaaggtctgtcgtgaa 1765
DB 1658 ACCTCTGTAGCTGATCGGACACAGCTCAACTTCAGACAGCGGACGGCTCCCTGGCA 1717
OY 1766 atatacagtgaaagctcagcttattatgtatgagagagatccaaagcaatgcccagtaa 1825
DB 1718 ACCTTGCTGTGCGAGTGCAGTACATGACAGCGAGAGACCCAGCAGGCTGTGCCGGTCA 1777
OY 1826 tctttgtaaatcttagctgtgttcagaatttccaagaaagccttaccagccgtatatactc 1885
DB 1778 TCTTTGGCAGAGCCAGCTGACAGTGAATTACCGGAGGCGCTTACACACCGGTGCTTACC 1837
OY 1886 ataaagctctccctgatttcatgaaatcaaggttaagcttccctgcttacttaactg 1945
DB 1838 AATAAAGTCCCGCCGAGTTTCACGAGGATTCAAGCTGCATCTTCCACCCGCTGACAG 1897
OY 1946 accacatcactgtcttacttacttacttacttactgttagtgcacaacaaacaaactatc 2005
DB 1898 AGAACCATCACCTGCTGTTCACCTTCTACATGTCAAGTGCAGCTGCCAGCCCGCGGACATG 1957
OY 2006 ctcttgaacacacagttgatatatacatalgataccaatgtctcagaatlgagcgttgaaga 2065
DB 1958 CCTGTGAGACACCCGTGGGCTTTACTTGATGCCACTGCTCAGACAGGGGCGCTGAGGA 2017
OY 2066 ctgagcaatttctgtcgcagatccatcttgaaaaaaccaacacaggttatactctgactgt 2125
DB 2018 CCGGCGCCCTTGTGCTCCACATGTCTGTGACACAGCCCGCCCGCAGCATTTCCCTGTCCA 2077
OY 2126 ctcccgaggtctcctacactgagcatgaatgggtlagataatcacaaaggtgttttlaatg 2185
DB 2078 CACCCGATGTGGCGCTTCCGGGCAAGCGCTGGGTGGAGCGGTACAAAGGCGCTGTTCAGTG 2137
OY 2186 ttgaagtgtgtgtgttctgcatacatalacaaagatcccttactcttgcacaatttttg 2245
DB 2138 TGGAGCTCACAGCCGTGCTCTGTGACACCCCAAGAGACCCCTACTGGACAAATTTCTTCA 2197
OY 2246 ctctgttcaatgctctgatatgaacaccgttcccaagtcgcaattggggacatcgtgaatca 2305
DB 2198 CCTGTGCTGACGCTCTGAGGAGGAGGCGCTCCATTCCTGAGGACACTGTGCTGA 2257
OY 2306 tggaaataacttgaataatgaatgaagacagatatttccagcaactgaattaccacagc 2365
DB 2258 GCGAGGGAACGCTGAGAGAGAGGCTGCGGGCCAGTCTTGACAGCACTGGCGCTGGCCAGCC 2317

OY 2366 tggaaacagtggtccgatttcttctcatctctctgtagataaactgatactttagttatata 2425
DB 2318 CCGAACCCCTTGTGGCTTCTCCACACAGCTGTGAGCAAGCTGTGCTGTGCTATCA 2377
OY 2426 gacctctgcatctgtgcgcaaatagtttaacctagtgtaaacacatcttltgaagcaatg 2485
DB 2378 GGGCCCCGATCATCACTAGTGGCAGATTGTGAACCTGGGCGGTGGAGCTTTGAAGCAATGG 2437
OY 2486 calcaattlaaaltcgaacttcaacaaaacttgaagaagaatcatalgcagacgtgcaaga 2545
DB 2438 CCCATGATGACACCTGTTTACACCGGAGCGCTGGAGGACCCAGATGCCCCGGGTCACT 2497
OY 2546 acagccttcttgacatalatataltatglttccgcctaaccaatcctaactcctaatt 2605
DB 2498 GCCACAGACTGGCTGCTGCTGCTACCTAGCGCTTTGCTTCTGGCCTGAGCAAGCCCA--- 2554
OY 2606 calcatcacaggtctctggggttttggggagatcagttgattatgccaatgagcttagat 2665
DB 2555 -----GCTCCCGAGTGGGGCCCCCTCCAGTGACAGTGAAGCTGACACATGGCCCGTG 2608
OY 2666 ctgcytgagaccttgcaagccttaatttaattcgtctcgaagccttgaataagcaatc 2725
DB 2609 GCTGTGTCGCCCGCAGAGCTTACTGGCGGTTCCAAAGCATCACAGACAGCAACC 2668
OY 2726 cagatatalctgggaactcccaagtcacagatgataaggttcgatcaatcatcogygagta 2785
DB 2669 CTGACCTGCGCGCTGCGCCCTGCTGTGAGATGACAGGTTTCCGCACTGCGCACAG-- 2726
OY 2786 aggtttagatgcttccaattccttgggtttaaacaactgtgttccaagaagctgcccatgg 2845
DB 2727 ----- 2726
OY 2846 gatccaaacccagttccaagtgacagaatcaacaagcgtatgataagttglaatcgtat 2905
DB 2727 ----- 2726
OY 2906 tgtctcgacaacagagacgtcaagtttcttaccaaatcaacygaacttaccacacta 2965
DB 2727 ----- 2726
OY 2966 aaaaagcttccaagggagctgtgcttgcagtggtgttggcagtgacaggttccgg 3025
DB 2727 -CAAGCTGTTCAAGAGAGACTGCTGTGAGTGGGTGTCACACACTGCGTACGCG 2785
OY 3026 aatcagcttgcacaagaagcgtgtcttcttgaattaaatgaagaagcagatggtgcacc 3085
DB 2786 AGGCCATCCCTCCAGCAGCGCTGTCTTCTTCCAGCTCATGTGAAGATATAGGCCCTGC 2845
OY 3086 attatactttaaatactgaacttgaagctcccaagaaagtcglttccaagaagtttca 3145
DB 2846 ACCTGCTGTTGGCCAGCGACTGTGACACACACCCCGCAAGCTGGCGTCCCGGACGCTTCC 2905
OY 3146 tggatgacatctgacgtctctgtcagcagatgtgtgtatataagtttaagatttcaga 3205
DB 2906 TGGACGACATCACTGCTGTGTGGTCTGTGGGCTGTGAGGATCAACCCGCTGTCCACA 2965
OY 3206 aggaacaagaatgtgttgaagaactcaatacaagaagccttcaatcttccaatgactgt 3265
DB 2966 AGGATGTGGAGCTGGCGAGACACTCAACGCCAGCGCTGTCTTCTTCCATGTGACCTTC 3025
OY 3266 tgtctgtttagacaagaagattgtttttagaccttataaagtcctgtctataaacaaggtgt 3325
DB 3026 TGTCCCTGAGTGGACGGGGCTTGTCTTCAAGCTGTGCGGGCCCACTCAAGAGAGTGG 3085
OY 3326 ctccaagcttactactaaltcaacgaatcccaagtglttctgtgtccttggaagcctggaatttc 3385
DB 3086 CCAAGCGGCTTCCAGTGTCTCCCTTAATCCAGACACGCTGTGACCTGTGCGCATGGAATTCA 3145
OY 3386 taagaatcatctgacgtatcatalgagacactatgttcatcttaacttaactcagacttacta 3445
DB 3146 CCGGCATCTGTGACAGCAGACGACACTAGCTGACCTCAACCTGCCCTGCTGCCCTGT 3205
OY 3446 ctccaactgatalcatcaactctgttcttcttgcacaactcagagttcttgattt 3505

D	3206	CACCTCCAGCCTCGCCCTCCCTCTGTGCTCTCCACCACTCCGAGGCTCCACCTTCT	3265
Q	3506	ctacgaatgtaacaagaccacaagaattgtcaaatatgttgaattatccgtcttcgcgc	3565
D	3266	CCAGCCAAAGCCCGGACCCCAAGGTGACGACGATGTTGGAATGATGAGTGCACATTCCGCG	3325
Q	3566	aacgacatatttggcaagacttgttgaacaagctgtgcgtgcaatttlaagacctgtatg	3625
D	3326	AGCACTGACCTTCTACTCTGGGTCCCTGCTGACGAGACTGGCACTGGCCCTCCAACTGTAGG	3385
Q	3626	ctgaagagactgtttggtattgtataaagaatgcataatattgtatacaaatattactctca	3685
D	3366	CTGAAGGGGCAATCTCTGTGCAAGAAAGGCATAGTGTGTGACAGCCCTGCTATGTG	3445
Q	3686	gtcacgaactcagaccgctgtactctgtaccctcagataaagctcgaatggccatgttgt	3745
D	3446	GCCATGACACTGACCCCGGCTACGGCGGACGCACTGTGAAGGCTGCTGTGGCCGACGCTGT	3505
Q	3746	atctaccctctatgtgtataltcatlgtgaactgttaccagctcagctgtatattacaagaa	3805
D	3506	ACCTGCCCCTGCTTTCGATTTGCACGGGATACCTTGCCACGGCTGCATGACTTGTGTGAGG	3565
Q	3806	ctcaacatacaagaggaaagac-----caatttgtatgtgccactgtatattatgaatgt	3859
D	3566	GCCAGGTGAGGGGTCAAGAGACTGCGCTTAATGCTTGACTGACACACAAAGGCGGAAGGG	3625
Q	3860	agaagcgaagtatgtataagccagaccgtttgcattgtcgaatgcgaagacatcgtgcctc	3919
D	3626	ACATGTGGGGGTACCACTCAACCCCTCTGTGGCCATGCGCATTTGCTGTGTGGCCCTTAGCCC	3685
Q	3920	aactaacaagacctgtgcagtttccctccctcagctcaacgaagtgtgcgaagacaacatacct	3979
D	3686	CTGGCTCCCGGGCCAGCATCTCCAGGGGCCACCAAGCGCTTTCGCGACAGGCTGTGCC	3745
Q	3980	tttaagcagaatcaagtgtgaagcctttgtactgtctactcttgggttttcaaatgtag	4039
D	3746	TCTGTGCTGAGCTCAAGCCGGACCTTGTGCGCTGTGTGCTGTGGGTGTGAAAAACACCG	3805
Q	4040	atgaaacagttctcaagaagtgtgtttacaagatctctcagctgttcgaagctaaacgcat	4099
D	3806	AGCCGGCGCTCTCGTAGGGGCTGGGCCACTGACCTGACACTCCCCAGCTGGGAGGCTGTGT	3865
Q	4100	taagatcgtcttatctctgtgtgtctgtcttgaagatataaaggaaaaagtgtttgaac	4159
D	3866	TGGACTTCTGTACTTGTGCTGTGGCTGCTTTGATGATCAAGGGGAAAAAGGCGCTTGMAC	3925
Q	4160	gaatgaatagcttgaacctttaagaatatcaaaagacatgaagcgaaagcttgaagaagcta	4219
D	3926	GCATCAACAGGCTCAATTCATAAAATATCTGTGATATGAAGGCGCGCTTAAGAGAAAGCA	3985
Q	4220	ttcttggagacatagtggtccagcgcaagaatgtgtacggtcgagccgagagacagctcgaga	4279
D	3986	TTCTGGGATACATCGGAGACTGACAAAGAAATGTTTGGCGAAGTCTGT-----4033	4033
Q	4280	gaagcccatctgtgaagtgtccttgtgaagltcaagaanaatttgaagltggaagaaatata	4339
D	4034	-----AGAGGAGGCCGCTTTGGGAATCCAGAGAAAGTGCCTGGCGGGAAGAGCGTCA	4084
Q	4340	ctcactgtgcgtcaaaaacacagagaagcttgcacaaatcaagaagcagagatgtgaacgaag	4399
D	4086	CACACTGGAACCAACCTCAAGACGCGGTGTGACAGACCAAGATGAATGACACACGAGG	4144
Q	4400	caactgtatgtgaagaaactgtctacagaagaacaaactatatttgaatcattagaga	4459
D	4146	CCTTGTGTGAAGGAACTGTGCACACCGAGGCAAGCTTAGTGTCTGTGACACTGTGGGA	4204
Q	4460	tttgtgttcagaccgttctcttlaacgaatccaagaagagacatttctgtgtgaagtata	4519
D	4206	TCACTGTGTCACAGCGGTATGCTTTCAAGAAAGCCGGGAAGCGCTTGTGGGGCACTGCTGA	4264
Q	4520	aagtgtactacacagaatgtgcgttgaaccaaagltgcaaglttatctatacaacatgttttg	4579

```

QY 5657 tggatgaagatcaaacagactcaatccgtgtagacaagtgtaaatagatcctaacaagg 5716
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5405 TCGTTGAGATTATCAAGAGACTTACCCCTGTGACAAAGTCCAGGCTTGACTCAAAAGG 5464
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 5717 catatctcagatcaactatgtgagccatacttgacacatatgagatgaagacagaa 5776
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5465 CCGTCAATCCAGATACCGTATGTGAAACCGCTTGTGATACCTAGAGCTCAAGACCGGG 5524
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 5777 tcacctatctcgacaaaataatacattctgtgatctcagctactgaaccccttactt 5836
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5525 TGACCTTCTTTGACCGCACTATGCGCTTGACATTCCTGTTCTGACGCCCTTGACGC 5584
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 5837 tagatggcgtgcccattgaggaacttcataacatcaaatcaaaagaagacattctacta 5896
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5585 CGGATGGGGCGGCACAGGGGAGTGCAGACACACAAAGCGTAAAGAGCCTGCTCAGCA 5644
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 5887 cgtctcatgctcttcttataataaacaagggtcaatgctactcataaagaagatca 5956
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5645 CCGACACAGCCTTCCCTTACATCAAGACTCGCATCCCTGTGTCACCGGGAGAGACGG 5704
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 5957 tcttaacacacatgaagtgtgctatltgagacatgcagaaaaagacagagatlgagcat 6016
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5705 TGCTGACGCCAGTGTGAGAGTGGCCATCGAGACATGCAAGAAAGACACGGGAGCTGGCT 5764
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6017 ttgacaacatcagatccgcagaccccaaaatgcttcagatggtactccagagatctg 6076
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5765 TTGGCACCGACAGACAGCCACACAGATGCTTAAGATGCTNACAGATGCTTCAAGGGCTCTG 5824
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6077 taagacacacagtgatcaacagggcccttggaagtgtcccaaggtttctcgtgaatac 6136
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5825 TAGGGCCACCGTGAACCAAGGCTCCCTGAGAGTGCGGCGAGGTGTTTGTAGCAGAGATCC 5884
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6137 ctatgacacacacacacgtcttcagacatcataataacacgcacgtctgctttaaagattta 6196
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5885 CGGAAGACCCCAAGCTCTTCGCGATCACAACAAATTGGCGCTCTGCTTCAAGGACTTCT 5944
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6197 ctaaaagtggtgaagatgcttcaagaanaataaagacttaattggccggttcaaaaag 6256
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5945 GCAAGAAATGTGAGAGATGCGCTCGGAAATAATAGGCCCTGATGGCGCGACAGAGAGG 6004
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6257 agtatacaaggaattg--ggaaactatcttgcgtcttaagaagagccctcagacctaga 6314
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6005 AGTACCAACCGTGAAGAGGAGGCACTACTGCGCTCGGAGAGGCTCTGACAGCCCTTGC 6064
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6315 tcacagaagaatgcctcagatataccaagccagatgtgttgcctccctccacagagattcc 6374
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6065 TTACCCAGCGCCTTCCCGACGCTGATGACACCCACCC-----ACCGGCTCTGAGAACTCC 6120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6375 ttccagtcgaatgagcttcgcaaatgtatctctaaac 6412
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6121 TTGAACAGAGCAAGTTTCCGAAAGCAGACCTCTGAGC 6158
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 10
AK055401      2607 bp      mRNA      linear      PRI 31-OCT-2001
LOCUS      Homo sapiens cDNA FLJ30839 fis, clone FEBRA2002429.
DEFINITION      AK055401
ACCESSION      AK055401.1 GI:16550120
VERSION      AK055401.1 GI:16550120
KEYWORDS      oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens fetus brain cDNA to mRNA, clone_1lb:FEBRA2
      clone:FEBRA2002429.
ORGANISM      Homo sapiens
      Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE
AUTHORS      Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,
      Otsuki,T., Sato,H., Makamatsu,A., Ishii,S., Yamamoto,T., Isono,Y.,
      Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
      Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
      Wagaatsuma,M., Morikawa,K., Kanehori,K., Takahashi-Fujii,A.,
      Oshima,A., Suzuki,Y., Sugano,S., Negahari,K., Masuno,Y., Nagai,K.

```

```

TITLE      and Isogai,T.
JOURNAL      NEDO human cDNA sequencing project
REFERENCE      2 (bases 1 to 2607)
AUTHORS      Isogai,T., Otsuki,T. and Sugiyama,T.
TITLE      Direct Submission
JOURNAL      Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
      Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
      (E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
      NEDO human cDNA sequencing project supported by Ministry of
      Economy, Trade and Industry of Japan; cDNA full insert sequencing:
      Research Association for Biotechnology (RAB); cDNA library
      construction: Helix Research Institute (HRI) (supported by Japan
      Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
      HRI, and Biotechnology Center, National Institute of Technology and
      Evaluation; clone selection for full insert sequencing: RAB and
      HRI.
COMMENT
FEATURES
      source      Location/Qualifiers
          1..2607
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="FEBRA2002429"
              /tissue_type="brain"
              /clone_lib="FEBRA2"
              /dev_stage="fetus"
              /note="Cloning vector: pME18SFLJ3"
              35..1933
              /codon_start=1
              /protein_id="BAB70917.1"
              /db_xref="GI:16550121"
              /translation="MAERRAPAKISRTVAEVRKQISQYSGSPOLLKMLNIVGNIS
              HHTVPLTEADPDVLEDTLTHPLAVDSGLRDLIEPPDDIEVVSPPDCRIVSA
              VPEEEMPHVDCIRSYEDMAIVIRYHKLIGFNPDLIDPKQEKQGLPKOVFES
              DEAPGNSYQDODDKIRMSYIDTPGSMAGSIPDKNSTPALLPILDRPNEE
              IDRDNDKQSNRKEIFALPSPDEEPRIELRSLVPDPKHEGQRLVLCISKEI
              IEPIEASLALDYVERKKEIKISENFYFINSQMGRLRHPVPAITTLARSAIFSIT
              KYPSQVFLVILKEVLQDGEAEPEPMIEFKEDATKNEKLEKLSQADQFORIG
              KYRMPFAMTALHMLIVSAGSLERDSIEVISTGERGSSWERNSSIVGRSLERT
              LKSDACNLTSFERPATLVTFNFPQEDGRSLDELYFLAMDERRSSVLRLRITQY
              LKDISAPENPHYCITPELLOVTLFPOEDRSRPRREILPEAPRVVYVNTYRLITQY
              YPQSLNFANRGSANRITVKVQFMIGEDPSNAMPSPSCTHIPVRLPLMCHTISFS
              LLSWDFIT"
BASE COUNT      843 a      500 c      506 g      758 t
ORIGIN
Query Match      26.7%; Score 1821; DB 9; Length 2607;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1821; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ggtccgttcgcccagacacatggtccgagggccggttcgcccacaaagatgagcaga 60
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 14 gttccggttcgcccagacacatggtccgagggccggttcgcccacaaagatgagcaga 73
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 acgtgtgacgcccgaagttagaagcagatctccggagacaatatagtggtctccccaactg 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 74 AGCGTGGCAGCGCAATTTGAAAGCAGATCTCGGACATATAGTGTTCTCCCAACG 133
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 ctcaaaaaccttaatatgttggaatatatccatcacacacacagtgccctcagaa 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 134 CTCAAAACCTTAATATGTTGTGGCAATATATCCATCACACACAGTGCCTTACCGAA 193
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 gcaatgagatccagtgatgtgaagaatataccatcattacacatccttggtgtgaattc 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 194 GCAGTGATCCAGTGGATTTGGAAGATTACCTCATTCATCTTGTGCTGTGATTTCT 253
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 gggccttaccgagatttgtaattcctccagatgatgatgaagtgtttatagtcct 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 254 GGGCCTTTACGGGATTTGATTGATTTCTCCAGATGATATGAAGTTGTTATATGTCCT 313
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 cggagctgcgaactctgttctcagctgtcactgaagaagtgaatggaatccacatgtt 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 314 CGGAGTCGAGAACTTGTTCACGCTGACCTGAGAAAGTGAATGATCCACATGTT 373
Qy 361 agagactgtataagaagttatatacagaagacttggcaattgtcatcaaaaaatcatataa 420
Db 374 AAGAGCTGTATAGAAAGTTATACAGAGACTGGCAATTGTCATCGAAAATATCATAAA 433
Qy 421 ttgggaacagaatttaatcccaatacatagaataacagaagaagaagcaaaagtttg 480
Db 434 TTGGGAACAGAGATTAATCCCAATATAGTAAACAGAAAGCAAAAGCAAAAGTTTG 493
Qy 481 ccaaaaaagaattttgaatctgaatgaagctcagaatggcaacagctaccagatgataa 540
Db 494 CCAAAAACAAGTTTGAATCTGATGAAGCTCCAGATGGCAACAGCTACAGAGATGATCAA 553
Qy 541 gatgaccttaaaaaagcgttcaatgtcacaatagaatgatacccaagaaggtgagctgct 600
Db 554 GATGACCTTTAAAAGAGCTTCAATGTCATAGATGATACCCCAAGGGGAGCTGGGCTTGT 613
Qy 601 agtattcttgacttgaanaaatcacttccctgatgtcttgcctcccaatttacttgatca 660
Db 614 ACTATCTTTGACTGAAAAAATTCACCTGCTGATGCTTTGCTTCCCAATTTACTTGATGA 673
Qy 661 actccaaatgaagaataagaccgttcaagaatgaatgaccaaaggaatcaaacgttcaaa 720
Db 674 ACTCCAAATGAAGAATGACCGTCAGATGATGACCAAGGAATCAAAACCGTCACAAA 733
Qy 721 gaacttttgcttgcattccatcaccagatgagaagaacaaatagaacggcttagtgtt 780
Db 734 GAACCTTTTGTCTTCATTCATCACCAAGATGAGAGAAACCAATGAAAGGCTTATGCTT 793
Qy 781 cctgaataacccaagaacatttggccaagaactctgttaaaatgttatacactcaag 840
Db 794 CCTGATATACCCAAAGAACATTTTGTCAAGACTTCTGTAATAATGCTTATCACTCAAG 853
Qy 841 ttggaatgtgaattgaacccaatttgcgaagtttgcttataatgatgtcaaggaagaag 900
Db 854 TTTGAAATGAAATGAACCCATTTTTCAGATTGCTTATATGATGTCAGAGGAAG 913
Qy 901 aaaaagattcagaacactttattttgaccttaatttgcagatgagaaggaaggttgtta 960
Db 914 AAAAAAGATTTTCAGAAAACCTTTTATTGTACCTTAACTTGAGAGATGAAAGGCTGTGA 973
Qy 961 cgtcacatgttaacacctgtgtgcattactaccctgtgcaagaatcagcaattttctatc 1020
Db 974 CCTCCACATGTCACACCTGCTGCCATFACCTACCTGCGCAAGATCAGCAATTTTTCATC 1033
Qy 1021 actatccttcccaagaatgttttctgttaataaagctagaagaagctctcagaagaaga 1080
Db 1034 ACTTATCCTTCCCAAGATGTTTCTTGTAAAGCTTAAAGAAAGTCTTACAGCAAGGA 1093
Qy 1081 gacatttgagagtgctgcagaacacataatgaatttcaagaagcagatgtcaccaagaat 1140
Db 1094 GACATTGGAGAGTGTGCAGAACCATATATGATTTTCAAGAAAGCAGATGCCACCAAGAT 1153
Qy 1141 aaagaanaacttggagaanaactgaagaagtcagaagatcagtttttgcacaagaacttggaa 1200
Db 1154 AAAGAAAACTGAGAGAAACTGAGAGTCAAGAGATCAGTTTGGCAAAAGACTTGGGAAA 1213
Qy 1201 tatgcatgaccttggcttgcagtcgaatcatttaagaatattgttaagcagtgctggg 1260
Db 1214 TATCCCATGCTTTTGTGCTTGAATCCATTTAAATGAATATTGTTAGCAGGCTGGG 1273
Qy 1261 agttggaagaagattctacagaagtagaatacagttacttgaagaagcaaaaggtcttgg 1320
Db 1274 AGTTTGGAAAAGAGATTCTACAGAGTGAATAATCAGTACTGGAGAAACGAAAAGGCTTTGG 1333
Qy 1321 tcaagaagaggaagaaacttactagatttggcagacagatcacttgaagaagacaagaagtgga 1380
Db 1334 TCAGAGAGAGAGAAATTTAGTATGTTGGCAGAGATCATTGGAAGGACAAAGGGA 1393
Qy 1381 gatgtagtcttgaacttgaagagcttgcagacagctactcacaagtgacaatattttt 1440
Db 1394 GATGATGCTTGTAACTTGACGAGGCTTTCGACCACTACTCTCACAGTGACAAATTTT 1453

Qy 1441 aagcaggaagaagaccgcttaagtgatgaagatctctacaagaattctctgtgatgata 1500
Db 1454 AAGCAGAGAGAGACCGCTTAAAGTGAAGAGATCTTACAAATTCCTGTGATATGAGA 1513
Qy 1501 aggcacatctgtcttctacggcgaactaagaacatttaccagctcagctcagaatagaact 1560
Db 1514 AGGCATCTCTTGCTTACGGCGACTAAGACATTAACAGCTCAGCTCAAGATAGACATT 1573
Qy 1561 tctccgcacctgaataatcccatlaltgcctaacctccggagctgtcctaagtgaagctt 1620
Db 1574 TCTCCCGACACTGAATAATCCCATTAATGCTTAACCTCCGAGCTGCTTAAGGAAGCTT 1633
Qy 1621 taccctgacagtagaagttagaaccttaccagaagaatcttgaagtttccgcaaggaagtt 1680
Db 1634 TACCCTGACAGTAGAGTTAAACCTACCAGAAATCTTAAGATTTCCCGAAGGAGATTT 1693
Qy 1681 tatgttccaaacactaactlacaagaatctctctacataatacctcagaagcttaattt 1740
Db 1694 TATGTTCAAAACACTACTTACAGAAATCTTCTACATATACCTCAGAGTGTATTTT 1753
Qy 1741 gccaatcgtcaaggttctgtcagaataatacagaagtgaaagctcagttatgtatgagag 1800
Db 1754 GCCAATGCTCAAGGTTCTGCTAGAAATATTAACAGTGAAGCTCAGTTTATGATGAGAG 1813
Qy 1801 gatccaaagaatcccatgagc 1821
Db 1814 GATCCAAAGCAATGCAATGCGG 1834

RESULT 11
AX174569
LOCUS AX174569 7215 bp DNA linear PAT 04-JUL-2001
DEFINITION Sequence 1 from Patent WO0142296.
ACCESSION AX174569
VERSION AX174569.1 GI:14598200
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Lu, P., Garman, J.D. and Candia, A.F.
TITLE Clasp-5 transmembrane protein
JOURNAL Patent: WO 0142296-A 1 14-JUN-2001;
Arbor Vita Corporation (US)
FEATURES
source
1..7215
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="full length human CLASP-5 cDNA"
112..6159
/note="human CLASP-5"
/codon_start=1
/protein_id="CAC43733.1"
/db_xref="GI:14598201"
/translation="MTHTNSLDVOLAQFLGPTDDLDLVFTPKRCPTLOPSLEEGV
EADPHVRCQVQYTRRWLVNRKNQSGPEIGCFKTKGRKQPHKTLQVTESELTLC
SEPAQAGRHNLNVLDVSGKPVYACQFDLSLPDRLENLQVSAEDEFKQNE
ARTNRQALEFALYPSVDEDAVEIRPVEPCKEHLKIRLVKLTLLKFELEIRPLA
STALDYQAEKRKISENFHCDLNSDQKFEFLAHNPVSAASSQAASVAVSYPSSDIY
LVVLTLEKVLQDGDIGCAEPYTVIKESGGSKSEIKLKIQAESFCORLGRKYMPPA
WAPISLSEFNSTLERVTVDVSGSPGPERTTLAOSRSLSERLSLENGVGN
FTSTLSVSSPFKQSGDRSDEDLKFLADYKRSSLDORVKSIPGLRLBSTAPEL
INCCILPEMLPVKPPENTRPHKEILFPTREVYVPHTVRNLLYYPPOLNVNKL
ASARNITIKIOPMGEDASNAMPVIFGSSGPEFLQEVYVATVYHNKSPDREVKIK
LEAKLTNHNHLLFTYHNSCOQKQASVETLLGSMELTLNERLQGSYCLPALEK
LEPNVSMASAEKVLPLONPPIKWAEGKGFVNIEOVAVSVTQDNHLEKFEFTLSLE
SQVTPRIVLDOKTISEMLHEIKSTICLASSREPLVFLHLVLDLKPOLSVOPW
TAGOTANFSQRFESVVAIANSKSLSDQHRNCLSLAYHYHFRDPEYORDV
KSGAPTALDPSHYHTGRTSAAAVSSKLLQARWSSNPDLQAGTHSADEVENIKS
SKTADRNCSRMYSYCGSSDAPSSPAARPAKSKHFEHLALQWVSYGMVKSMAOHV
HMDKRDSPFRTRFSDRPMDDITITIVNVVTEISALALVKKPKKENEOAEKNIISLAFLL

YDLISLMDRGFEVNLIRHYCSQISAKLSNLPILLISMRLEFRLICSHENYLNINLFEM
NADTAPTPSPCISQSSSSSSSFODKISMFDLVSEYQOQHFLEGLFETLALD
AEGGISKVORKAIVSAHLSLSSHDLPKREPVKLVIAKLVIPVYIIDLPLQIC
DPTVADRRRTSGSDSEBGAGAINONVALIAGNENPKTSGIYLSISPKOYVML
NADPTRLMILCFIMKKNADOSLIRKRIADLPSTQJNRILDLFIVLYLEPKYKQSS
DKVSTVOLKSRVKAKALEALLRGEAREMRRRAPDRPGINELRWKKEQTH
WRQANERKIDTKALEIDQEPALISGNLATEALILIDQENIQAQSLADKSDLSGLV
RVLVNSINCDSQSTTVLTHCFALRALIAGDGLFEEVEOCDFLOHVIHSCSSMD
VTRSOACATLYLIMRFESFNARVKYMDVMSIASIVGRAPFMEHLRSLFTLL
AYSBEDTAMQMTPEPPOVEELLINLSIYDVKMRBFQDPEMDLMDLIRAKSYOA
SPDLRLTWNMAEKHKKCTEAMCMVIAHAIYAEIISHEIDSYLPYGSVQN
ISSNVEESVYSDELTPDEDVACQYFESGLVGLDEAALFSTGILYEVNIV
KLVIPLIAHREFRKLTLSKLRADFSLVNDKRMETGYRVEGFSKEDLDEO
EFVYKEPATIKLEIPISHRLEAFYQOCFGAEVIVIDSPVDKIDLPKAYIOITFV
EPYDEEMKDRATYEFKRNENLRREMYTPTELEGRGEGHEOYRNNYLTTHMFP
YIKTRISIOKEEFVLPIEVALIEDMKKLTLOLAVAINOEPAPKALOWILOSVAAT
VNOGPLEVAOVFLAELPADPKLIRHNKRLKLCREKIMRGGEVKNRKLITIDOREY
QOELKKNYNLKENLRPMLEKRIPELKYPLFRVESOKRDSFHRSSFRCETOLDSGS"
BASE COUNT 2005 a 1734 c 1721 g 1755 t
ORIGIN

Query Match 25.6%; Score 1749; DB 6; Length 7215;
Best Local Similarity 58.2%; Pred. No. 0;
Matches 3613; Conservative 0; Mismatches 2360; Indels 238; Gaps 21;
Qy 173 ttaccgaagcagtagatccagtgatgttgagaagattaccattactcatccttgctg 232
Db 71 TTTAGACCCCTGTGAGCAGTGAACCTTTCTATACACACCTGAAAGCC 130
Qy 233 tggattcctggccttaccggagatttgatgaatttcctccagatgatatgaattgtt 292
Db 131 TGGATGTGACACTTCCGAGAGCTCGGGACTTCACTGATGACGACTTGACGTGTG 190
Qy 293 atagtccctgggactgcgaactcttcttcaactgtgacctggaagaag---tgaatgg 349
Db 191 TCACGCCAAAGAGATGATGAGACTTTCAGCCCTTTGCGGAGAAAGGGGTTAACTGG 250
Qy 350 atccacatgttagagactgataagaagttatagagaagacttggagcattgttcacga 409
Db 251 ACCCTCATGTCAAGGAGCTGTTCAGACCTTACATCCGATGGCTAATCTGTGAACCGGA 310
Qy 410 aatatacataattgggaacaggaatttaaccataacatagataaacaagaagaagc 469
Db 311 AAAACCAAGAGATGCCAGAAATGCTGTAAAGAACTGATTCGAAAGATTTC 370
Qy 470 aaaaagtttgcacaaaagaattttgaatctgatgaagctccagatggcaacgatacc 529
Db 371 ACAAGACGCTTCCGAAACAGACGTTTGAGTGGGA-----AACCTTGGAGTGCAGTGAC 424
Qy 530 aggatgatacagatgaccttaaaagacttcaatgtaacatagatgatacccccaagggt 589
Db 425 CCGCTGCTCAGGAGGCCCCCGCCACTTAAAGTGCTGTGGACGCTGTCTGGGAAAGGCC 484
Qy 590 gctgggactgtagtaacttgaactgaaaaatcaacttcctgactgtcttgcctcaatt 649
Db 485 CCGTCACTGCTGTGACTTTGACCTCCGACGCTGACGCTGCAAGAGGCGCTAGAAAAAC 544
Qy 650 tactgtatcgaaactcaaatagaagaatagacgctcagaatgatagacaaagaagaatca 709
Db 545 TCCTTCACACCAAGTGAAGTCCGAGGACTTTCAGAAAGCAAGAGAGAGGCCCGCGAGACCA 604
Qy 710 accgtcacaaagaacttttgctgtgacatcacacagatgagaagaagaacaataagac 769
Db 605 ACAGGACAGGCGAGCTCTTTGCGCTTTTACCATCAGTGGAGAGAGAGATCTCTGGAAA 664
Qy 770 ggcctagtgcttcgataaccacaaagaacttctgacaaagcttctgtaaatgct 829
Db 665 TACGTCCAGTACAGAAATGTCACAGAAACACTGGGCAACAGATATTTGTCAAGTTGC 724
Qy 830 tatacctaagttgaaattgaatgaaccatttttgcgaagtttgcttataatg 889
Db 735 TGACCTTGAAGTTGAGATTAATTTGAGCCCTGTGTCGACGATTCCTCTACGATG 784

Qy 890 tcaaggaataaagaagatttcagaataacttattttgaccttaattcctgagcagatga 949
Db 785 TTAAAGAAAGAAAGAAAGATCTCGAAAAATTTTCACTGTACTCGAAGCTGACAGTTCA 844
Qy 950 aagggttgtaagtcacatgtaaccactgctgcattactacacctggaagaatcagaa 1009
Db 845 AAGGATTTTGGCAGGCTCACAGCCCTTTCAGTGGCGCATCAAGTGAAGGAGATGGCAG 904
Qy 1010 tttttctacacttctctcccaagatgcttttctgttaataagctcagaanaagctc 1069
Db 905 TCTTCTCAGTCACTTCCCTGCTCTCAGACATCTACTGATGATCAAGATTGAAAAAGTCC 964
Qy 1070 taagaagaagacattggaagatgtgcagaaacatataatgatttccaaagaagaagatg 1129
Db 965 TGGAGCAGGAGGATATTGAGACTGTGCGAGAGCCCTACACGGTTATCAAGAAAGATGATG 1024
Qy 1130 ccaccaagaataaagaanaaacttggagaacatgaaagatcagaagatcagtttgc 1189
Db 1025 GTGAAAAGAGTAAAGAAAGATGAAAAACCTAAACCTCAAGCTGAATCTCTGCCAGC 1084
Qy 1190 gacttggaaatatcgacagccttctgctgtgactgacatccatttaagaatatgta 1249
Db 1085 GTTTGGGGAATACCGGATGCTTGTGCTGGGACCCATTAAGCTTATCAACCTCTT-- 1142
Qy 1250 gcaatcgtggagatttggagaagatttccagaagtagaanaatcagttatggaagaa 1309
Db 1143 -CATATGCTCCACCTCTGAGAGAGGAGGTAACTGATGTGA-----CTGTGCTGTTGGGA 1195
Qy 1310 aagggtcttgatcagaagaagaatctcagatgttgatggcagacatcacttgaaga 1369
Db 1196 GAAGCCAGAGGCTGAACGAGGACATTTGGCCCAATCTGAAAGGCTTTC---TGAAGAG 1252
Qy 1370 caacaagtgaagatgactgttgaacttgaagagacttgcacagactactcacaagta 1429
Db 1253 CCCTCTCTTGGAGAAAGGGGTTGGATCCAACTTCAAAACCTTCACCTGAGCGTTA 1312
Qy 1430 caaatttttaaagcagaagaagagacgcttaagatgataatctcacaattccctg 1489
Db 1313 GCAGCTTTTCAAGCAGAAAGAGATCGCTTACGATGAAGACTTATTCAGATTTTTG 1372
Qy 1490 ctgatatgaaagagcactcttctgcttaagcagactaagaacctatcagctcagctca 1549
Db 1373 CTGACTTAACAAAGATCATCATCTTACAGAGACGATCAAGTCAATTCAGAGCTGCTTA 1432
Qy 1550 agatagaactttcccgcaactcgaataatcccatattgcttaactccggagctgttc 1609
Db 1433 GACTGAGATTTTACAGCTCCAGAGATCATCAATTGCTGTCTGACTCTGAAATGCTGC 1492
Qy 1610 aagtgaagcttaccctgacagtagagtttagacctcagaagaatcttagagttcccg 1669
Db 1493 CCGTGAACCCCTTCTCTGAAACCGGACACAGCCCGCACAAGAGATTTTGGATTTCCAA 1552
Qy 1670 caaggaatgttaatgttccaacactactacagaataatctctctacatalaaccctcaga 1729
Db 1613 GGCTGAACCTTTGTAAACAAACCTAGCATCAGCCCGGAGAACTTAAATAAAGATCCAGTTTA 1672
Qy 1790 tctatagagatcccaagaatgcagatgcggttaactctgttgaacttagctgttcag 1849
Db 1673 TGTGTGAGAGAGATGCTAGCATGCGATGCGGTCTATCTTTGAAAAATTCACGCGGCTTG 1732
Qy 1850 aatttcaagaagcctatatacagcgttagtatataatacaggtctcctgtattcatg 1909
Db 1733 AATTTCGAGGAAGTGTACACAGCTGTGTATACATATACATATAGTCTCTGACTTTTANG 1792
Qy 1910 aagaatacgaagtttaagcttccgtcacttaacttaactgacatcatcactgtcttact 1969
Db 1793 AAGAAGTGAATAATTAAGCTCCCGCTTAAGCTACAGTAAATCACACGCTCCTGTTACCT 1852

Db 3848 TACTTTTCATGTGTGTTATGTTTGAATTAAGGAAAAAGAGTTTCGACAAAGTCA 3907
Qy 4166 atagcttgaccttaagaatccaagaacatgagagcaaaagcttgaagaactatcttg 4225
Db 3908 GTACCAAGTCTCTGACAGAGCAAGGATGTCAAGGCCCGGCTTGAAAGAGCTTTGCTGC 3967
Qy 4226 ggaagcatagtgccaagcaagaatagtbacggcgaagccgaagacagctcgaagaagcc 4285
Db 3968 GTGGGGAAGGGGCGAGAGGAGATGATGCCCGCGGCTCCAGGAAACGACCGATTTC 4027
Qy 4286 catctggaagtgaccttggaaagtaagaataattgagtggaagaagaatagatgactact 4345
Db 4028 CA-----GGCCTTAATGAATAATTTTGAATGGAAGAAAGACGACACATTT 4072
Qy 4346 ggcgtcaaaacacagagaagcttgacaatatcaagaagcagaattgaaacaacgaactga 4405
Db 4073 GGGGCAAGCTAATGAGAGAGTACGATAAACAAGGCCGATTGATGATCAAGAACCTTTGA 4132
Qy 4406 ttgatggaacctggctacagaagaacaaactaatcatcttagatatacttagagattgtg 4465
Db 4133 TCAGTGGCAATCTGGCTACAGAACACATTTAATCATCTTGATATGCAGGAAAAACATTA 4192
Qy 4466 ttcaagacgcttctgttaacggaatccaagagagcattctgtgtggaagtgttaaaatgc 4525
Db 4193 TCCAGGCGAGCTCGGCTGTGACTGTAAAGACGCTGTGGAGGTGTTGAGGTGC 4252
Qy 4526 tactacacagcatgacctgtataccaagaatgacattatctacaacactgtttgtctaac 4585
Db 4253 TGTGTAATTTCTGMACTGTATCAGANTACCACTTACTACTGCTGCTTTCACAAAC 4312
Qy 4586 agagaagccttggttcaaaagttccctgaactcttatttgaagaagaagacagagcagtgtg 4645
Db 4313 TCCGTCCTCATCCCAAGTTTGGAGACTTACTTTCGAAAGAGAGGATGMAAGAGTGT 4372
Qy 4646 ctgattatgctcaaggtcttcgcacactgtagcaagtagatggtatacaatagtcac 4705
Db 4373 TCGACTTAATGTACCAAGTCTCTGCAACCACTGACAGCAGCATGATGTCCACCGGAGCC 4432
Qy 4706 acccagagtcctcccttactactactaactgaagcaaaacttggagatttggataacttg 4765
Db 4433 AAGCTGTGCCACCTTTTACTCCTCATGAGTGTCAAGTTTGGAGCCACCGTAAATTTTG 4452
Qy 4766 ccaaggttaaatgcaagtgaccaatgtcaatactcctctgtgtggcaactctcagaatt 4825
Db 4493 CAAGAGTGAAGATGCAATGAACATGCTCCATGTCATTTGGTGGAAAGCACACAGACT 4552
Qy 4826 ttaagaagaattcttaagaagttctctaaagactatattgacatattgtctgaagaatc 4885
Db 4553 TTAATGAAGAGCACCTGGAAGATCTCTTGAGGACAAATTTTGCCCTAATTCAGAAAGAGCA 4612
Qy 4886 tgaatttgaggaaacaacattctctgtacaggtccagatctggttttcaatctccata 4945
Db 4613 CAGGCATGCAAGTGACTCTTTTCCCAACCGAGTGAAGAACTTCTCTGATATCTGAATA 4672
Qy 4946 tgaattctcttgatactgttgaataatgaagaacacacagagagatctcgttaattgtatg 5005
Db 4673 GCATCTTAATATACACATGAAATGAAGGAAATTCAGGAAGATCTCAGATGCTTATAGG 4732
Qy 5006 acctaatgttaagaatttgcgaaggttctacagactctccaga---ggagattgacagtg 5062
Db 4733 ATCTATATGACGAATTTCCAAAGTATCCAGGCAATCTCCGATCTGCGGGCTGCTGGC 4792
Qy 5063 tgcagaacatgcaagcaagcactcagaagaagcaatcatgcttgaagctcagagtttc 5122
Db 4793 TCCAGAACTGCGCAAGAAACACCAAGAAAGTGTCTACACGAGAGCTGCTCATGTGCC 4852
Qy 5123 tagtccactcagcagcaacttgttctgaatatattgagcatgctggaagccgaagaatc 5182
Db 4853 TGGTCAAGCCCGCTGCTAGTGTGCTGAGTATCTGAGCATCTGAGAGACACAGCTACC 4912
Qy 5183 ttccgttgtagatglttaacattcgaatatattcactaaatggttttagaagaatctggcg 5242
Db 4913 TTCCCGTGGGCAGTGTCAAGCTTCAGAAATATTTCTTCCAATGTGCTGAGGAGTCTGTGG 4972

Qy 5243 tctcatatgatgtggtatctccaagatgaagaaggtatctgctctggaataacttactg 5302
Db 4973 TCTCTGAGGACACCTGTGCACCTGACGAGAGATGGGTGGCCAGGCCAGTACTTCACCG 5032
Qy 5303 agtcaggaacttgttggtattacttgaagaacagcagctgtcttctctctatagtctgcatg 5362
Db 5033 AGATGGCCCTGTGAGGCTCTCTGGAGCAGCGCGAGGCTCTTCAGACAGGGAGGCTTAT 5092
Qy 5363 atgaagcatttaagaagtttcaagaagttactatctctatctcatatgaagatcagagtg 5422
Db 5093 ATGAGACAGTTAATGAGTGTCTACAACTGTGATCCCATCTTAAGAAAGCATTCAGAAAT 5152
Qy 5423 caaagaactatccaacatctcgttgaacttgaagaacttgaagaagcacaatgttcatc 5482
Db 5153 TCCGGAAGCTGACACTACACTACAGCAAGCTGAGAGAGCTTTCGACACATCGTTAAACA 5212
Qy 5483 agagtaactgctggagggagatgttggcaactatttctggttggtttttagaaca 5542
Db 5213 AGGATCAT----AAGAAATGTTTGAACCTACTTCGAGTGGATTCCTTTGGATTCGA 5266
Qy 5543 agtccgggaatttggatlgaaagaatttggtttacaagagcctgcaataaccaactg 5602
Db 5267 AATTTGGGATTTGGATGGAACAGAGATTGTCTACAAAGCCTGCAATTTACCAACTTC 5326
Qy 5603 cagagatattccaagatttgaaggaatttcaaggaagaagatttggagaagatgtgtgtg 5662
Db 5327 CTGAGATCTCATACTAGACTGAGAGCAATTTATGTCATATGTTTGGTGCAATTTGTGG 5386
Qy 5663 aagtaatacaagaacttaactccgtgtagaacaagtgttaatatagatcctaacaagcatata 5722
Db 5387 AAGTGAATTAAGACTCCACTCCGTGAGCAAAACCAAGTTGGATCTTAACAAGGCTTACA 5446
Qy 5723 ttcaagttaccctatgtggagccacttcttgacaacatagatgaagaagaagaataacct 5782
Db 5447 TACAGATCACTTTTGTGGAGCCCTACTTTGATGAGTATGAGATGAAGAAGCAGGCTCACAT 5506
Qy 5783 attcgacaaaattacaatcttcgtcgaattcagtgactgtatacaaccttactttagatg 5842
Db 5507 ACTTTGGAAGAAATTTCAACCTCCGGAGGTTCAATGATACACACCCGTTACACCTGAGG 5566
Qy 5843 gccgtgcccatggggaaactctatgaacaattcaaaagaagaacattctgactagctcc 5902
Db 5567 GGGGCGCTCCGGGAGAGCTGTCATGACAGTACGAAAGGAACACAGTCCAGCCACTATGC 5626
Qy 5903 atgccttctcttatataaaacaaggttcaatgttcaactataaagaagaatcatcttaa 5962
Db 5627 ACGCTTCCCTTACATCAACAGACAGATACGCTCATCCAGAAAGGAGATTGTGTTGA 5686
Qy 5963 caccaattgaagttgtctatltgaagacatgcagaanaaagacacagagatttggatctgcaa 6022
Db 5687 CACGATTTGAAGTTGCCATTGGAAGACATGMAAGAAAGAACCTCGAGTTAAGCATGCCA 5746
Qy 6023 caactcagaatcccgcaagcccaaaaatgcttcaagatgtgatactcagaagatctgtatgca 6082
Db 5747 TTAACCCAGAGCCCGCTGATGCAAAAGATGCTTCAAGTGGCTGCAAGGCTGTGGGAG 5806
Qy 6083 ccaacagtgaatcaggggaccttggaaagtgcgaagtttcttctgtctgaataactagtg 6142
Db 5807 CTACTGTAATCAAGGACCACTGGAAGTAAACCAGTGTTTTGGCTGAATTCCTGCTG 5866
Qy 6143 acccaagctcttcaagacataataaactgcagacttgcgtttaaagaatttactataaa 6202
Db 5867 ATCCAAAACTCTATGCACATCAACACAGAATGTGAGTTATGCTTTAAGGAATTCATCATGA 5926
Qy 6203 ggtgtgaagatgccttaagaanaaataagaagcttaattgtggccggttcaaaagagatc 6262
Db 5927 GATGTGTGAAGCTGTGAAGAAAACAAAGCTGTCTACGCGCAGCACAGAGGAATATC 5986
Qy 6263 aaaggaatctgggga--aaatatcttcgctttaaagaagccctacagccctagatc--ac 6318
Db 5987 AGCAGGACTCAAAAAGACTATTAACAAGCTTAAGAAGAACTTCAGGCCCAATGATGAGC 6046

[illegible]

D	b	1958	CCGTGAGAACACACCGGTGGGCTTTACTTGTTGATTCGCCACTGTCGTGACACACAGGGCGCTGAAGGA	2017
Q	y	2066	cttgccagtttctgttcagctgcagctcattctggaaaaaaccacaggtcattctgtactgt	2125
D	b	2018	CCGGCCCTTCTGTCTCCCACTGTCTGTGGAGACACGCGCCGCCCAAGCTATTTCCCTGTCTCA	2077
Q	y	2126	cttcctgaagttcccttacccttgcatgaaatggtgtagataataacacaagaagtggttttaatg	2185
D	b	2078	CACCCGATGTGGCGCTTCCGGGCATGCGCTGGGTGGAGCGGTACACAAAGGGCTGTTCAGTG	2137
Q	y	2186	ttgaagttgttgtgttttcgtctcatalcaatacaagaatcccttactcttgcanaatttttg	2245
D	b	2138	TGGACCTCACGGCGCTGTCTCTGTGGACACCCCCAGAGACCCCTACTGTGACAAATTTCTTCA	2197
Q	y	2246	ctctgtgtcaatgctcttgatgaaacaccctggttcccaagttccgaattggtggagactggaatca	2305
D	b	2198	CCCTTGTCCACTCCCTGGAGAGGAGGAGCGCTTTCCATTTCCGCTCAAGGACACTGTGCTGA	2257
Q	y	2306	tggaaataacttagaanaatgtaatggaagacagatcttcagactgtaactatcccgac	2365
D	b	2258	GGGAGGGCACTGTGAGGACGAGAGCTGGGGGCAAGTCTTGACGACACTGGCGCTGGCGAGCC	2317
Q	y	2366	tggaaocagtggtgccgaatttcttactctctgtctagataaaactgtaacttttagttatta	2425
D	b	2318	CCGAACCCCTTGTGGCTCTTCCACCCACAGCTGGTAGAACACTGCTGGCTGTGGTCATCA	2377
Q	y	2426	gaccctccgtgacttgctgtggcaaaatgafttaactgaagccaagcatcttltgaagccatgg	2485
D	b	2378	GGCCCCCGATCATCACTGAGTGCCAGATTGTGAACCTTGGCCGTGGAGCCTTTGAAGCATGG	2437
Q	y	2486	catcaatataaatactgacttccaaaanaacttggaaagaaatcatatgacagatgycagaa	2545
D	b	2438	CCCATGTACTCAGCCTTGTCTTACCCGAGACCTTGGAGCGCACCCAGATCCCGCGGTCACT	2497
Q	y	2546	aagacctcttgcatacatataatcatatgttttccgcctacccaatcacttaacctaat	2605
D	b	2498	GCCCAAGTGGCTCTCTCACTGTCACCTAGCGCCTTTCGCTTCCTGGCACTGAGCCA---	2554
Q	y	2606	catcatcaacagtcactcctgtygggttttgggaagaaatcgatgcattatgtccacaatgtagat	2665
D	b	2555	-----GCCCTCCGGATGGGGGCCCTCCAGTACAGTGGAGGCTCCACACTGGCCCGTG	2608
Q	y	2666	ctgctgtgtagacctgtcaagccttaatttaaatcgttctcogaagccttgtaatagcatc	2725
D	b	2609	GCTCTGTGTCGCCCGCAACCCCTCTACCTGTGGCGCGTTCACAAAGGACTACGACACACCAACC	2668
Q	y	2726	cagatatatctgtagactcccaacgctacacagatcgtatgaaagtgcataatcctggaatla	2785
D	b	2669	CTGACCTGCCCTGGGCCCTTGCGCTGTGTGATGTAGAGAGTTCCCGCATCTCGGCCAG--	2726
Q	y	2786	agggtttagatcgtctccaaatctctcgtgtthaacactgtgtgtccaaaagctgccacatggg	2845
D	b	2727	-----	2726
Q	y	2906	tgtcttcgacacagagacgttaagtlttlticaaacaattaaacygagcgttacaactha	2965
D	b	2727	-----	2726
Q	y	2966	aaaagcttlttcaagagagagctgtgcttttgcattggttgttgcagtgycagcgttggg	3025
D	b	2727	-CAAGTGTCTCACGAGGAGGTGGCTCTGTCAATGTGGTGTCAAGCAATGCGGTACGGG	2785
Q	y	3026	aatcagcttctgacaagaagcctgttcttcttlttgaattatggtlaaagacatgctgtgac	3085
D	b	2786	AGGCAATCTCTCAAGCAGCGCTGTGTTCTTTCCAGCTANTGTGTGAAGATATAGGGCGTGC	2845
Q	y	3086	attataacttaatgataaacttgagactccaaggaanaagtcgttttccagaagtcca	3145
D	b	2846	ACCTGTGCTTGGCCACGACACTAGACACACCCCGCAAGCTGTGGCTTCCCGGACGCTTTC	2905

[illegible]

QY	4220	tttttggggagcatagttccagagcaagaatatgttgcgcgaagccgagacagctcgaga	4279
Db	3986	tttttggggtaccatccgagagctcgacaaataatggttggggcaactctcgagagc-----	4038
QY	4280	gaagcccatctgaagtgcctcttggaagtcgaagaatatctggatgtaggaagataatga	4339
Db	4039	-----accctgtttgggaattccagagaaactgccttgccggaaagacgtca	4084
QY	4340	ctcactgagctcaaaaacacagagaagaacttgcacatatcaagacagagatltgacagaa	4399
Db	4085	caacactggaagcaaaacctcagacacgcgctgagacaagacaagcagaatgtaattgaaacagag	4144
QY	4400	caactgttgaatggaaacccgtgcctacagaagcaaacctaatacatlttagatacaatagaga	4459
Db	4145	cccttggtggaagggaacctggcgaacccgaagcagacctactgtgttcggacacacttgaga	4204
QY	4460	ttgtttcagaacggtttctgttaaggaatccaaagaagacatctctgtgtgaagtcata	4519
Db	4205	tcattctgtagagacggatgatgcttttcagaaagcccggaagacgcttttggggcagtgctcta	4264
QY	4520	aagtgctactacaacagcagctgctgttaaccaaagltgcaatttatacacaacgtgtttg	4579
Db	4265	aggtttgctgtaacacacttgggagatggccagagtgctcctcttcttgcagcattggccttg	4324
QY	4580	ctacaagaagagccttggtttcaagaatttctctgaactcttatttgaagaagagacagc	4639
Db	4325	ccaccacagaggccctttgtgtccaaattcccggagctgctgttccgagagacacgcagac	4384
QY	4640	agtgctcatattatbcccctcagagctcttcgcgaacatgtagcagtagcatcgttaacatc	4699
Db	4385	tgtgtgccgacctgttcccttgaggtccttctacgacactgttgccagccgcatcacaccattcc	4444
QY	4700	ggtcaaccccaagtgctccctcttaacctatactatgaggaanaacttgaatttggaaata	4759
Db	4445	gcacgcacccacagccgctcgctgtaactctgcatggagacagaacttcgattcgggcacaca	4504
QY	4760	actttgcagaggtttaaatgcaagtaaccaatgtcactatctcctctgttgggcacatctc	4819
Db	4505	acctttgcccgttggaagatgcaggtcacacattctcttctgctccctggggggagcagccg	4564
QY	4820	agaattttaagaagaattcttaagaactctcctaagaactctattgaatatgtgtgaag	4879
Db	4565	aaactctcagtaaaagacaccttgcacagcttcaactaaaaacatcctacacttaattgttgag	4624
QY	4880	aagatctggaatttgaaggaaacaacatcttcctgcatcaggttccagagctcgtgtttcaatc	4939
Db	4625	acgacatgagggctgcggggacacacaccttgcgacagacagstccagacacctgaatttcaacc	4684
QY	4940	tccatatgatcttctctatctatcttgaanaatgaagaacaaccaagagaatccttgaatgt	4999
Db	4685	tgcacatatttcttgacggcagcagctggaagatgtaaaagaaacaccagagagacaccttgagatgc	4744
QY	5000	tgaattgacttaatgtacagaagaatttgcacaaggtttaccagaacctctcaga---ggagatga	5056
Db	4745	tcattgacctctattgtacgaataatttgcocgggcttaccagagaccttcacggacacttgcgctga	4804
QY	5057	ctctgtttcagaacaatggcagcagaacactcagaacgaagcaatcatctgaagttgaac	5116
Db	4805	cctggttgagaaactatggccgggaagcagacggagacttggcgaacacacgccagagccgcc	4864
QY	5117	agtgctctagttccactcagcagc	5138
Db	4865	agttgatggtgacccggccgc	4886

RESULT	13
AK024436	
LOCUS	AK024436
DEFINITION	Homo sapiens mRNA for FLJ00026 protein, partial cds.
ACCESSION	AK024436
VERSION	AK024436.1
KEYWORDS	GI:10440380
SOURCE	His (Full Insert sequence).
	Homo sapiens adult spleen cDNA to mRNA, clone:as00026.
	4577 bp mRNA linear PRI 29-SEP-2000

4149 agtcttgaacgaatgaatagcttgaaccttaagaatcaaaacatgagacaagct 4208
1276 GAGTTGTGACAAAGTAGTACCAAGTCTCTGACGAAGTAAAGGATGTCAAGCCCGGCT 1335
4209 tgaagaagctatcttggagcatagtgycagagcaagaatlgtaagcgaagcgaag 4268
1336 GGAAAGGCTTTGCTCCGTGGGGAAGGGGCCAGAGGGGAGATGATGCCGCCGGGCTCC 1395
4269 aacagccgaggaagcccactcgtgaagtgcccttggaaagtcagaagaatttgagtgag 4328
1396 AGGGAAAGCAGCCGATTTCCAGGCCCTAA-----ATGAAAATTTGAGATGGA 1440
4329 gaaaatagaactgaactgagctcaaaaacagagaagcttgaacaatcaagaagcagat 4388
1441 GAAAGGACGACACATTTGGCGCAAGCTATAGAAAGCTAGTAAACAAAGCCGAGTT 1500
4389 tgaacaggaagcaactgattgataaaccctgctacagaagaacaaacctaatcttaga 4448
1501 AGATCAAGAAAGCCTTGATCAGTGGCAATCGCTACAGAACACATTTAATCATCTGGA 1560
4449 tacattagagatgtgtctcagacogtctctgtaacggaatccaaagagagatctctgg 4508
1561 TATGCAAGAAACATTATCCAGGCGAGCTGGCTCTGACTGTAAAGACAGCTGCTGGG 1620
4509 tggagtgctaaaagtctactacaagcagcagctgtaaccagtgcaagtgctatctaca 4568
1621 AGGTGTTCTGAGGTCCTGTAATTTCTGTAAGTGTGATCAGATACCCTACTGAGC 1680
4569 acaactgttctgcaacagagagccttggttcaaaagttccctgaactcttatttgaaga 4628
1681 TCACCTCTTTGCAACACTCGGTCTCATGCCAAGTTGGAGCACTTACTCTTGAGAA 1740
4629 agagaagagcagtggtgagcttattgctcagcgtctccgacactgtagagtagat 4688
1741 GGAAGTGGAAAGTGTTCGACCTATGTACCAAGTCTGTGACCACTGACAGAGCAT 1800
4689 cggtaacaatagctcacaccagtgcccttacccttaactaagaagcaaaacttga 4748
1801 GGAATGCACCGGAGGCAAGCTGTGCGCACCTTTACCTCCCTCAGATGAGGTTTGG 1860
4749 gattgggaataacttgcagaggttaaaatgcaaglaccaatlgtaactatccctctgt 4808
1861 AGCCACCACTAATTTTGGCAAGAGTAAAGATGCAATTAACATGCTCCCTGAGCATCTTGGT 1920
4809 gggacatctcgaattttaagaagaattcttaagaogtctcctcaagacatactgac 4868
1921 GGAAGAGCACAGACTTTAATGAAGAGCACTGTGAAGATCTTGAGAGACAATTTTGGC 1980
4869 atatgtgaagaagatctgaaatagggaaacaacatctcctgtaacagtgccagatct 4928
1981 CATATCAGAAAGAGACACAGCATGCAAGTGCATCTTTTCCACCACAGTGTGAGACT 2040
4929 ggcttcaactccataatgattctctctgtaactgtaagaatgaaagaaacacaggaaga 4988
2041 TCTCTGATCTGATAGATCTTATATGACACAGTGAAGATGAGAGATTTTCAAGAGA 2100
4989 tccctgaagatgtgatgatactaaatgtaagaatctcaagaggttcagagactccaga 5048
2101 TCTTGAGATGCTTATGATCTCATGTACGAAATGTCGAAGATACAGGCACTCTCGTA 2160
5049 ---gcatgtacactgtgtgcaagaacatgycagagcaagcactcagaagaacaaatcagc 5105
2161 TCTGCGGCTGACCTGGCTCCAGACATGGCAGAGAAACACACCAAGAAAGTGTCTAC 2220
5106 tgaagctgcaagtgctcagtcacactcagcagcactgtgtcgtgaatttgagcatgct 5165
2221 GGAAGGCTGCAATGTGCTGTGTGTCACGCCCTGCTTAGTGCTGATCTGAGCATCT 2280
5166 ggaagacggaataatctctcgtgtgagatgtaacatttgaagaatttatactcaagt 5225
2281 GGAGGACACACCTACCTGCCCCGTGGGAGTGTACAGTTCCAGAAATTTCTTCCAAAGT 2340

5226 tttagaagaatctgcggtctcagaatgagtgtgtaacctccagatgaaagaagtactcgtc 5285
2341 GCTGGAGAGTCTGTGCTCTGTGAGACACCTCTGACCTGACGAGAGATGGGCTGTGCGC 2400
5286 tgaanaaacttactcgaagtcagacttggagattacatgaaacaagcagctcctcct 5345
2401 AGGCGAGTACTTACCCAGAGTGGCTGTGAGGCTCTGAGACAGGCGCCGAGGCTTT 2460
5346 ctctatgctgcgaatgataagcaagttaagttaagaatgaatgaatgaatgaatgaat 5405
2461 CAGCAGCGGAGCTTATATGAGACATTAATGAGCTGTACAACTGGCTATCCCATCTT 2520
5406 tgaagcctaatcggagatgcaagaacatccaaactcattcgaatgaatgaagaact 5465
2521 AGAAGCGCATGCAATTTCCGAGAGTGCACATCTCACAGCAAGCTGCAAGAGCTT 2580
5466 caagcaaatgttcaatcaagtlactggtcgtggagcgaagtgttggcaactatctcgtgt 5525
2581 CGACAGCATGTTTAAACAGATCAT-----AAGAGAAATTTTGGAACTTACCTTCCGAGT 2634
5526 tggtttattgaagcaacaggttcgggagtttggatgaaacaagaattglttaacaagagcc 5585
2635 TGGTTCTTTGGATCCAAATTTGGGATTTGGATGAACAGAAATTTGTTCAAAAGAGC 2694
5586 tgcataaccaaaacttgaagataltcaacagatlgagaggaattttagcgaagaagatt 5645
2695 TGCAATTAACCAAGTTCCTGATGATCTCAGATAGACTAGAGGATTTTATGTGTAATGTT 2754
5646 tgaagagagtggttgaaglaalcaagaactcaactcgtatgacagagtgtaaataga 5705
2755 TGTGTCAGAAATTTGTGAAAGATTAAGACTCACCTCCTGTGACAAACCAAGTTTGA 2814
5706 tcccaaaagagcatatltgaattactatgtggagagccacttattgacacataagat 5765
2815 TCTTAACAAGGCTTATTAAGATCACTTTTGTGGACCTCTACTTGTGATGATTAAGAT 2874
5766 gaagagacagaalcaaccatttgcacaaaactcaactcgtcgtatgtaactgtaactgac 5825
2875 GAAAGACAGGTGTCATACTTTGAGAAAGATTTCAACCTCGGAGGTTCATGTACACAC 2934
5826 acccttactttagatggtgcgtgcgaatggggaactcaatgaaactcaaaagaagac 5885
2935 CCCGTTACCTGAGGCGGCGCTCGGGAGAGCTCAAGACAGTCAAGAGGAACAC 2994
5886 catctgaactagctcatgctcttcccttalttaaaacaaggttaagtcactata 5945
2995 AGTCTGACCACTATGACAGCTTCCCTACATCAAGACCAAGATGAGCTCATCCAGAA 3054
5946 agaagagatcacttlaacacaaatlgagatgtcatlgagagacatgagaaagaacaca 6005
3055 GGAAGAGTTTGTGTAACACGATTTGAAGTGGCATTTGAAGATGAAGAAAGAACCCCT 3114
6006 ggaagtgcatttgcacacatcagaatcccgagaccccaaaatgcttcaagatgact 6065
3115 GCACTTGACAGTTGCCATTTAAACAGAGCGCCCTGATGCAAAATGCTTCAGATGTGCT 3174
6066 ccaagagatcgtgaagcacaagatgaatgaagggcccttgaaggttggccaggttct 6125
3175 GCAAGGCTCTGTGGACTTACTGTAAATCAGGAGCACTGTGAAGTGAAGCCCAACTGTTTTT 3234
6126 gttcgaataactagtgtaaccaaagctcttcaagacatcaataaactcgaactcgtc 6185
3235 GCGTGAATTTCTGTGATCCAAACCTTATGCAATCATCAACAAAGTGAAGTATGCTT 3294
6186 taaagatttactaaaagtggtgaagatgcttcaagaanaaataagagcttaattgggcc 6245
3295 TAAAGAAATTCATCATGAGATGTGTGAGAGCTGTGAAGAAACAAACGCTCTCATCAGCGC 3354
6246 gttcnaaagagatcaaaaggaattgggga--aaactactcgcgttaagaagccct 6303
3355 AGACCAAGGGAATATAGCAGGAACTCAAAAAGAACTATTAACAGCTAAAGAGAACT 3414
6304 aagccctatgac--acagaagatccctcagttatccaaagccagatatt 6349

Db 3415 CAGGCCAATGATCGAGCGAAATTCAGAACTGTACAGCAATATT 3462

RESULT 14
LOCUS BC008335 2825 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, clone IMAGE:3506202, mRNA, partial cds.
ACCESSION BC008335
VERSION BC008335.1 GI:14249910
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2825)
Strausberg, R.
Direct Submission
Submitted (25-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanina Chiu, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McAraay, Steven
Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Scott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

REFERENCE AUTHORS TITLE JOURNAL

REMARK COMMENT

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNLN at: <http://image.lnl.nih.gov>
Series: IRAL Plate: 23 Row: d Column: 12.

FEATURES Source

CDS

1. 2825
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3506202"
/tissue_type="Kidney, renal cell adenocarcinoma"
/clone_id="NH_MGC_14"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
<1..2585
/codon_start=3
/product="Unknown (protein for IMAGE:3506202)"
/protein_id="AAH08335.1"
/db_xref="GI:14249911"
/translation="ORSLASMDSDTEGEGDIATNPVSAMAIAGPLPAGSRASI
SOGPPTRAGCALSAESSRTLLACVAMVKNPEPILQAWADLLPOLGRLLDILY
LCIAAFYKGGKAKFERINSLTFKKSLSMDKRLKEALLIGTGAQEVRRSREPSGN
PENVRKSTYTHNKQTSRDYDKTDEDEHAEALVEGNATLADSLVADTEITVQVLM
SEARSLVGLAVLVLSLGSASALELQGLATQRLVSKFPELLFEDETECADLC
LRLRHGSGRISTIRTHASASLYLMMONTEIGHNARVQVMTSLSSIVGTQNS
EHLRSLKTLITYAEDMGLRSTFAEYQODLHMFLHTVTVMEKEOEPENLI
DLMRIARGXGSPDLRLTLMONMAGHAEHGAEEAOCMVAAALVAYEALALDHD
RHLPGCVFQNSNYLSESAISDILSDEGPGSGKFTGLVGLLEOAGVPT
MGGLYEAVNEVKNLPILEAHDYKLLAVHKKLQDAFTKIMHSGSMERVCTFR
VGRGAHFGDLDEOEFYKEPSTIKLAELSHLEEFYTERGDDVEILLKSDMPVKS
KLSQKAYIQITTYEPFDYELKDRVTEYDNRVGRTEFLCTPPTPDGSAHDELPED
HKRRTLSYDTHAPYIKTRIRVCHRETVLTPVEVAIEDMKRELAETAEDEPDPA
KMLQVLOGSGVGTQVNGPLEVAOVFAELREDEPKLFRHNNKRLCFKDECKCEAL
RKNNALIGPDQKEYHRELEENYCRLEALQPLLTQRLPOLMAPTPPGLNSLNRAFR
KAL"

BASE COUNT 641 a 835 c 827 g 522 t
ORIGIN

Query Match 16.4%; Score 1119.4; DB 9; Length 2825;
Best Local Similarity 66.0%; Pred. No. 16-250;
Matches 1694; Conservative 0; Mismatches 841; Indels 30; Gaps 4;

Qy 385 gaagtgcagcggaagtatgaagaacgacgcttgccatgcaatgcaggaatcg 3912
Db 48 GAAGGGGACATTCGGGTACATCAACCCCTGTGTGGCATGGCCATTCCTGTGGCCCC 107
Qy 3913 gtcccttaactaaacaagcgcttgcaagttctccctcaacgcaagatgcagcaac 3972
Db 108 CTAGGCCCTTGCTCCCGGGCCACATCTCCAGGGCCACAAAGGCTCTCGCGAGGC 167
Qy 3973 actaaccttcaagaagaatgaagtcgaagcctttgatctgtctacttggttctcaa 4032
Db 168 TGTGCTCTCTCTGTGATGTCMAAGCCGACCTTGTGGCTGTGTGTGGCTGTGAMA 227
Qy 4033 aatgcagatgaacagttctcaagaagttgtttaacagatctcagttcagcttaac 4092
Db 228 AACACGAGCCGGCGCTCTGACAGCCCTGGCCATGTGACCTGCCAGCTGGGA 287
Qy 4093 cggcctatagatctgtcttactctctgtgtgtctgtctgtgtgtgtgtgtgtgtgt 4152
Db 288 CGTCTGTGTGACCTGTGATGTCMAAGCCGACCTTGTGGCTGTGTGTGGCTGTGAMA 347
Qy 4153 ttggaacgaatgaatgctgtgacctttaagaanaatacaagaatgaagcaaaacttga 4212
Db 348 TTGGAACGATCAACAGCCCTCACATTCMAAAATCTGTGATGMAAGCGCGGTAGAG 407
Qy 4213 gaagcattcttgaggcagtagtgccagcaagaatgtgctggcgaagccgagagacg 4272
Db 408 GAAGCATTTGTGGGTACATCGACGCTGCACAGAAATGTTGCGGAAAGTGTGAGAG 467
Qy 4273 ctgcagagaagcccatctggaagtgctcttggaagtcgaagaatattgagtggaaga 4332
Db 468 -----AGCCGTTGGGATCCAGAGAACGTCGCGTGGCGGAAG 506
Qy 4333 gatatactactcggcggttaaaacacagagaagcttgcaaatcaagagcagaattga 4392
Db 507 AGCGTCACACACTGGGAAGCAAACTCAGACCGCGTGGAGAACCAAGATGAAATGGA 566
Qy 4393 cacaagacactgattgatgaaacctgcgtgtacagaagaacaaacttaacttaataa 4452
Db 567 CAGGAGGCTTGTGGAAGGAACCTGGCAACCGAGGCAAGCCTTGTGAGGACGA 626
Qy 4453 ttagaattgtgtcagaacgcttctgtlaacggaatccaaagaagacattctgtgtga 4512
Db 627 CTGAGATCATCTGCGACAGCGGTGATGCTTTCAGAAAGCCGGGAGAGCCCTTGGGGCA 686
Qy 4513 gtgtcaaatgtctactacaacagatgacctgaaccagaatgacgttacttaacaac 4572
Db 687 GTCTGAGGTTTGTCTGACACACCTGGGACAGGCCAAGTGCCCTCTTCTTGGAGCAT 746
Qy 4573 tgttgcctacacagagagcgtgtgttcaagaattcccgaaactctattttaagaag 4632
Db 747 GGCTGCGCACCCAGAGGCGCTTGTGTCCAAATTCGCCGAGCTGTGTCAGAGGGGAC 806
Qy 4633 acagaagcagtgctgattatgctcctaagcgtctctcgaacactgttagcagtagcatcgt 4692
Db 807 ACGGACGTGTCGCCAGCTGTCCTGAGGCTCTTACGACACTGTGGGACCGCGCAAC 866
Qy 4693 acgaatcgtcaaacccacgctcctccttactactactaagaagcaaaacttgaatt 4752
Db 867 ACCATCGGACGCGACGCGCGCTGTGCTACCTGTCTATCGAGAGAACTTCAGAGAC 926
Qy 4753 ggaataacttcccgaggttaaaatgcagagacaaatgactactccctcctgtgtggc 4812
Db 927 GGCACACACTTTCGCGTGTGAAGATGACAGGTACCATGTCTCTGCTGCTGCTGCTGCTG 986
Qy 4813 acatctagaattttaatgaagaattcttaagaagcttcttaagaactataattgacatat 4872

Db	987	ACGACGAGAACTTCACTGAAGAGCACTGGCAAGCTTCACTATAAACCATCTCAACCTAT	1046
Qy	4873	gctgaagaagatcttggaatctgaaggaaacaacatcttcctgatcaagtcagatcttggt	49322
Db	1047	GCTGAGGAGGACATGAGGGGCTGCGGGACACACACTTCGCGAGACAGTGTACAGACCTGATG	11060
Qy	4933	ttcaatctccatcttgatcttcttcttgatcttggaatgaagaacacacagagatcct	49922
Db	1107	TTCAACTGTCACTGATGATCTCGAGGACACGGTGAAGATTAAGAACACCAAGAGACCTT	1166
Qy	4993	gaaatgtctgatctgatctaatgtlcaagaatctgcgaaggtgtlaccagaccctccaga---	5049
Db	1167	GAGATGCTCATGCACTGATCATGTACAGATTGGCCGGGCTTACCAAGGGCTCAACCGGACTT	12266
Qy	5050	cgattgacccttggtttgcagaacatctgcgaagcagaccctcaagaacgaagcatatgctgaa	5109
Db	1227	CGGTGTACCTGGTTGGCGAACAATGGCCGGGAAGCACGCGGACTTGGGAACCAACGCCAG	1286
Qy	5110	gctgcagaagtctagtcacactcagacagcaactgtgttcgaatcttgagcatcttgag	5169
Db	1287	GCCGCCCAAGTGCATGATGGGCACGCGGCCGCCCTGTGTGCTGATTAAGTGGCTGTGCAG	1346
Qy	5170	gaccggaatatctctcctgttgggaagtgtlaaacatttcagaatatctcaatgttlla	5229
Db	1347	GACCACCGCCACCTGCCCCGTTGGCTGCTGTTCTTCCAAACATCTCATCCAAACGTGCTA	1406
Qy	5230	gaagaatctcgcgtctcaagatgtatgtgatatctccagatgaagaagratctgcctctga	5289
Db	1407	GAGGAGTCCGCCATCTCCGACGCATCTGTGCCCCGACGAGAGAGGCTTGTGCTCCGG	1466
Qy	5290	aaatacttaactcagatcagaagcacttgttgggaattactgaaacaagcagctgtctctctct	5349
Db	1467	AAGCACTTCACTGAGCTGGGGCTGTGAAGGTTGCTGGAACAGGACCGGCTACTTCAAC	1526
Qy	5350	atgctgtgcactgtatagaagagttatgaagtttcaaaagtctatcttccatctatcaga	5409
Db	1557	ATGGGCGGGCTCTACGAGGGCGTGAATGAGGTTACAAACCTCATCTCCACTCTGGAA	1586
Qy	5410	gctatctcggagatcgaagaacactatccaaatctcagtgtgaacttcaagaaagcatctcgc	5469
Db	1587	GCCCACTGTGATCAAGAAGGCTGGCCGGCGGTGCACGCAACCTGCAGAGAGGCTTCAAC	1646
Qy	5470	aaaattgttcatcagagttactgtgctctggagcggaatgtttggacctatlttcogtgtgt	5529
Db	1647	AAGATCATGACACAGAACTTCGGCTGGAGCGCGGTGTTGGGACGATTTTCCGCTGGGC	1706
Qy	5530	tttatatgaacaaagtctcgggagtttggagttgaacaagaattgttttcaagaagcctcga	5589
Db	1707	TTTCAAGCGGCCCACTTCCGCTGTGACCTGATGAGACAGAGATTGTGTCAAGAGGCAATCG	1766
Qy	5590	ataaccaactctgcagagatalctcaacagattggaggattttacgcgaaagaatttga	5649
Db	1767	ATTCACGAAGCTGGGAGAGATCTCACACCGGCTGGAGAGATTGTTCACGCGAGATTTGGC	1826
Qy	5650	gagagatctgttggaagaatcaaaagccttaatccctgttgaacaagtgttaaatgaatcct	5709
Db	1827	GACACAGCTGTTGAGATTATCAAGACTCTTAACCTGTGTGAACAGTCCAAAGCTTGACTCA	1866
Qy	5710	aacaagcatalatctcagatcttacctatgttggagccatcttggacaacatatgaagtga	5769
Db	1887	CAAAAGGCTTACATCCAGATCAAGTATGAGAACCTGACTTTGATTAACCTACAGACTCAAG	1946
Qy	5770	gacagaaatcactattcttcgcaaaaattacaatcttcgtcgtatcatgtactgtacaacc	5829
Db	1947	GACCGGGTGAATCTTTGACCCCAACTATGGGCTTGGCACTTCTGTTGTGACGCGCG	2006
Qy	5830	tttactttaaagtccgtgcgccatggggagactctatgaaacaattcaaaaggaagcaatt	5889
Db	2007	TTTCAAGCGGATGGGCGGCACACAGGGGAGACTGCCGAGAACACAAAGCTGTAAGCGCTG	2066
Qy	5890	ctgaactcgtctcatgtcccttctctatlatthaacaaggtcgaatgtcactatagaaga	5949

[illegible]

TYRVEFGSKFSDLDDEQFEVYKEPAITKLPISHLRLAEFYGGCFGAEEVETIKDSTP
VDRFKLDPNKAYIOITFEVEPYDEYEMKRVYFEENFMNIRRMYPTEPLEGRPGE
LHEOYRNNVLTTHMAFPYIKTRISVIOKEEFVLPTLEVAIDEMKTTLOLALINOE
PPAKMLQVLOGSVATVNOGPLEVAOVLAPIDAPDKYRHHNKLRLICPFEIIMRC
GEAVERKKRLITADOREXOELKNNKLKENLRPMIERKIPILYKPIIRVESOKRDS
FHSSFRKCTQLSQGS"

BASE COUNT 1161 a 898 c 957 g 1010 t
ORIGIN

Query Match 15.8%; Score 1079; DB 6; Length 4026;
Best Local Similarity 62.98; Pred. No. 2.8e-241;
Matches 1836; Conservative 0; Mismatches 1035; Indels 46; Gaps 9;

QY 3443 ttaccaccgcatccatcacctctctgttcttctgcacacatccagatctcgat 3502
DB 8 TGANTGCGTACAGTCTCCACATCTCTGTCTCTTCATATCTTCCAGACTCAAGCT 67
QY 3503 ttctacgaatgtaacaaagaatgcaaatatgttgaat--atccgtgcctt 3560
DB 68 CCTGCTCCAGCTTCAGAGACGAGATCGCACATGTCGATCGGACTTCCAGACTAC 127
QY 3561 ccgcacaaga-ttatttgcagacttggttaacagcgctgcgtcaattttagacc 3619
DB 128 CCGCAGCAGCACTTCTCCACCGGGGCTCTCTTCCAGAACTGGCTGCCCTGGATG 187
QY 3620 ctgactgtaagagactgttggattgataaagaatgcaatgataatgataatc 3679
DB 188 CCGAAGGGGAGAGATCGAGAACTACAAAGAAAGCTGACAGCAATTCACAGCTGC 247
QY 3680 tctccagtcagactcagaccgcgtaacttgcacctcagataaagctcgatggca 3739
DB 248 TAAGTTTCACAGACTGACCCACGCTGTGTCAACACAGAGTGAAGTCAAAATCGCCG 307
QY 3740 tttgttctacacctgattgattatcatgaaactgtaacctcagctgtatgttta 3799
DB 308 CCGTTTACCTTACTTGTGTTGGATCAATTTGGATGCTTTGCCACACTCTGTGACTTTA 367
QY 3800 cagaactcacatcaacgaagaagacaaattgatacgaactgactgattagaagt 3859
DB 368 CAGTTGGAGATACCTCGCAGATACGCGACACAGTGC-----TCGATGAAAGAACAG 418
QY 3860 agagcgaagtaataagcagaacgctgcacatgcacatgcagagacatcgctcc 3919
DB 419 AAGGAGCGGCGTCAATACCAAGATGTGCTCGCATAGCAGGGAATTAATTCAAT 478
QY 3920 aactacaagcgctgcagattctctctcactacgtaacgaatgagcagaacacta 3979
DB 479 TGAATAACAAGT-----GAATAGTGTCTTCTTCCCTATTAACCACTAACATGC 532
QY 3980 tttaagcagaatcaagtcgaagccttctgacttctgacttctgacttctcaaatg 4039
DB 533 TGAACGGGAGACACTACGCACTCATGATCTCTCTCTCTGATCATTAATAATCTG 592
QY 4040 atgaacaagtctcaagaagtgtttacagatctctcagctcttgcagctaaacgg 4099
DB 593 ATCGAGCCCTCATTAAGAGATGATGCTACCTGACATCAACGACTCAACAGATTT 652
QY 4100 tagatctgacttactctgtgtgtctgttgaatgaataaggaataaagtgttgaac 4159
DB 653 TAGATCTACTTATCTGTGTGTATGTTTGAATGAAGGGAATAACAAGTTCTGACA 712
QY 4160 gaatgaatgacttgaacctttaagaatcaaaagacatgagagcaaaactgaaagac 4219
DB 713 AAGTCATACCCAGAGTCTCGAGAACTCAAGGGATGTCAAGGCCGCCGCTGGAAGGCTT 772
QY 4220 ttcttggagagatagtgccaagcaaatgtaagcggaagcgaagcgaagctcgaga 4279
DB 773 TGTGCTGTGGGAGGGGAGGAGAGATGATGCGCCCGCGGCTCCAGGGAAGGAGCC 832
QY 4280 gaagccatctggaagtgcccttgaagtcagaataattgagtgaggaagatatatg 4339
DB 833 GATTTCGA-----GGCTTAATGAATAATTGAGATGAGAGAAAGACAGA 877

QY 4340 ctcaactggcgtccaaaacacagagaagcttgacaatatcaagcagagattgacagcgaag 4399
DB 878 CACATTGGCGGCGAGCTAATGAAAGCTAGATTAACAAGCCGAGTTAGATCAAGAAC 937
QY 4400 cactgattgataagaaactgctgacagaaacacactaacttataagataactagaga 4459
DB 938 CATTGATCAGTGCATCTGGCTACAGAGCAACATTTATATCTCTGATATGAGGAAA 997
QY 4460 ttgttcttaagacgcttctctgaacgaaatccaagagagacatcttggtagagctaa 4519
DB 998 ACATATTCAGAGGAGAGCTCGGCTGTGACTGTAAAGACAGCTGCTGGAGGATGTTCTGA 1057
QY 4520 aagtgctactaacagcagctggcgtgttaaccagaagcgagttatataaactgtttg 4579
DB 1058 GGTGCTGTGAAATTTCTGAACTGTGATCAAGATCAACCTACTGACTCACTGCTTTG 1117
QY 4580 ctacacagagagccttggttcaaaagttccctgaaacttatttgaagaagagacagc 4639
DB 1118 CAACACTCGGTGCTCATCGCAAGTTTGGAGACTTACTCTTGAAGAGAGAGTGGAAC 1177
QY 4640 agtgctgattatagcctcagagctctccgaactgtagcagtagcatcggtagaatac 4699
DB 1178 AGTGTTCAGACTTATGTCAACAACTCTGACCACTGACAGCAGATGATGTCAACC 1237
QY 4700 gttcacacccagtgctcccttactactaataaggaacacttgaatttggaaata 4759
DB 1238 GAGAGCCAGCCTGTGCGACCTTTTACTCTCTATGAGTTTCAGTTTGGAGCCACCACTGA 1297
QY 4760 acttggcagaggtttaaatagtcaggtacaaatgtcaatacctccttggtagacatc 4819
DB 1298 ATTGTCAAGAGTAAGATGATCAAGTACCATTCCCTGSCATCTTGGTGGGAAGAAGCAC 1357
QY 4820 agaatcttaagaagaatcttaagaagctctctaagaactatltgacatatgctgaag 4879
DB 1358 CAGACTTATATGAAGAGCACTGAGAAAGATCCTTGAGGCAATTTTGGCTATTCAGAG 1417
QY 4880 aagatctggaatgaggaagcaacatctcctgactcagatccagatctggtttcaatc 4939
DB 1418 AGGACACACCCATGACGATGACTCTTTTCCACCCAGGTGAGGAACTTCTGTATATC 1477
QY 4940 tccataatgattcttctgatactgtgaaatgaagaagcaacagagagatccctgaatgt 4999
DB 1478 TGAATATGATCTTATATGACACAGTGAATAATGAGGGAATTTGAGAAATCTGTAGATGC 1537
QY 5000 tgattgacttaatgtacgaatcttccaaggttaccagacctctccaga--gagatga 5056
DB 1538 TTATGATCTCATGTACAAATTTGCCAAGATTTACAGAGCATCTCTGATCTGCGGCTGA 1597
QY 5057 cctggttggagaacatgagcagaagcactcagaacgaacatcatgtgaagctgac 5116
DB 1598 CCGGCTCCAGAACATGCGAGAAACACACCAAGAGAGTGTACACGGAGGCTGCA 1657
QY 5117 agtgcctagtcacactcagaagcactgtgtcgtgaatatttgaacatgctgagagacgga 5176
DB 1658 TGTGCTGTGTGACGCGCGCTGCTGTAGTGTGAGTATCTGATGAGTGTGAGGAGCACCA 1717
QY 5177 aatatcttccgttggagatgtgaacatttcaaatattcatctaaagtgtttagaagaat 5236
DB 1718 GCTACTGTGCGGTGGGCACTGCTACCTTCCAGAAATTTCTTCCAAATGTGCTGGAGGAGT 1777
QY 5237 ctgcgctccagatgagtggtatctccagatgaagaagatgctgcctggaataatc 5296
DB 1778 CHTGTGCTCTAGAGACACCCCTGTCACTGACGAGAGATGGGTGTGCCAGGCCAGTACT 1837
QY 5297 ttactgagtcagagacttgtaggattacttgaacaaagcagctgtctctctctatgctgt 5356
DB 1838 TCACCGAGAGTGGCTGTGAGGCTCTCTGAGACAGCGCGAGAGCTCTTACAGACGGGAG 1897
QY 5357 gcatgataagaagcattgaatgaatttcaaaagtaactatcttcatatgaagtaatc 5416
DB 1898 GCTTATATGAGACAGTTATGAGTGTACAAAGTGTGATCCCTCAAGAGGCGATC 1957

OM of: US-09-737-246-2 to: GenEmbl:* out_format : pfs
Date: Jul 16, 2002 2:46 PM
About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODE=frame+pat.model -DEV=xlh
-O=/cgrn2.1/USPRO.spool/US09737246/runatc.16072002.075430.26464/app-query.fasta.1.2180
-DB=GenEmbl -OFMT=lastab -SUFFIX=p2n.rge -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -STARP=1 -MATRIX=blonsum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US09737246.@CGN1.1.6982 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPPX -WAIT -THREADS=1

Search information block:
Query: US-09-737-246-2
Query length: 2090
Database: GenEmbl:*
Database sequences: 1797656
Database length: 187333701
Search time (sec): 5569.590000

score_list:	Strid	Orig	zScore	EScore	len	Documentation
gb_pat:AX172874	+ 10815.00	12925.95	0.0	6828	1	AX172874 Sequence 1 from Pat
gb_pat:AX173175	+ 7136.50	8519.28	0.0	6372	1	AX173175 Sequence 1 from Pat
gb_pat:AB051558	+ 6436.50	7665.07	0.0	4311	1	AB051558 Homo sapiens mRNA for
gb_pat:AX172980	+ 6171.50	7367.86	0.0	4200	1	AX172980 Sequence 107 from Pat
gb_pat:AX172880	+ 6157.50	7351.24	0.0	4144	1	AX172880 Sequence 7 from Pat
gb_pat:AX172932	+ 6157.50	7351.24	0.0	4144	1	AX172932 Sequence 59 from Pat
gb_pat:AX174569	+ 5785.50	6889.11	0.0	7215	1	AX174569 Sequence 1 from Pat
gb_pat:AB037816	+ 5472.50	6528.59	0.0	4866	1	AB037816 Homo sapiens mRNA for
gb_ov:AY016023	+ 4986.50	5905.47	0.0	167215	1	AY016023 Spheroides nepheli
gb_hlg:AC017375	- 4700.00	5571.69	5.5e-302	73515	1	AC017375 Drosophila melanog
gb_in:AC008318	+ 4690.00	5550.25	8.6e-301	166626	1	AC008318 Drosophila melanog
gb_in:AC005950	+ 4690.00	5543.40	2.1e-300	301639	1	AE005950 Drosophila melanog
gb_pat:AX172978	+ 4626.50	5516.89	2.8e-302	3327	1	AK055805 Homo sapiens CDNA FL
gb_pat:AX172943	+ 4626.50	5519.63	4.3e-302	3253	1	AX172978 Sequence 105 from Pat
gb_pat:AK055493	+ 3528.50	4200.10	1.4e-225	4577	1	AK054436 Homo sapiens mRNA for
gb_pat:AC005436	+ 3235.50	3856.24	1.9e-206	2452	1	AK054933 Homo sapiens CDNA FL
gb_pat:AC008335	+ 3207.00	3820.46	1.9e-204	2825	1	BC008335 Homo sapiens, clone 1
gb_pat:AK055401	+ 3122.50	3720.14	7.4e-199	2607	1	AK055401 Homo sapiens, clone 1
gb_pat:AX174571	+ 3025.00	3598.30	4.5e-192	4026	1	AX174571 Sequence 3 from Pat
gb_pat:AX174664	+ 3025.00	3598.30	4.5e-192	4027	1	AX174664 Sequence 96 from Pat
gb_pat:AX173233	+ 2523.00	3039.41	6.1e-161	2149	1	AX173233 Sequence 59 from Pat
gb_pat:AX173022	+ 2480.50	2940.44	2.0e-155	6454	1	AX173022 Sequence 1 from Pat
gb_pat:AX255048	+ 2345.50	2776.95	2.6e-146	7506	1	AX255048 Sequence 7 from Pat
gb_in:BC021414	+ 2273.50	2706.55	2.1e-142	1900	1	BC021414 Mus musculus, clone 1
gb_pat:AX174662	+ 2245.50	2672.32	1.7e-140	2014	1	AX174662 Sequence 94 from Pat
gb_pat:BC019102	+ 2071.50	2459.62	1.2e-128	2904	1	BC019102 Homo sapiens, Stimul
gb_pat:CN000005	- 2063.00	2400.26	2.4e-125	205035	1	AL049870 Human chromosome 14
gb_pat:AB028981	+ 1929.50	2281.36	1.0e-118	5862	1	AB028981 Homo sapiens mRNA for
gb_in:CEMF46H5	- 1900.00	2224.16	1.6e-115	38886	1	U41543 Caenorhabditis elegans
gb_pat:CN001RGX	+ 1780.00	2065.59	1.1e-106	139887	1	AL160233 Human chromosome 14
gb_pat:AX173028	+ 1632.50	1928.84	4.4e-99	4391	1	AX173028 Sequence 7 from Pat
gb_pat:AX173118	+ 1613.50	1908.47	6.0e-98	4391	1	AX173118 Sequence 97 from Pat
gb_in:AC008354	- 1609.00	1859.20	3.4e-98	159350	1	AC008354 Drosophila melanog
gb_in:AX058220	- 1509.00	1851.67	6.8e-95	257224	1	AE003617 Drosophila melanog
gb_pat:AX173120	+ 1484.50	1754.23	4.8e-91	3899	1	AX058220 Sequence 90 from Pat
gb_pat:AX173120	+ 1484.50	1754.23	4.8e-91	3472	1	AX173120 Sequence 99 from Pat
gb_pat:AX056684	+ 1328.50	1568.43	5.3e-79	3152	1	AK056684 Homo sapiens CDNA FL
gb_pat:AX058192	+ 1183.00	1396.27	2.0e-69	2610	1	AX058192 Sequence 62 from Pat
gb_pat:AF465831	+ 1120.50	1311.97	1.0e-64	5896	1	AF465831 Arabidopsis thaliana
gb_pat:AF194407	+ 1111.00	1318.58	4.3e-65	1242	1	AF194407 Homo sapiens unknown

gb_pat:AK054649	+ 1097.00	1294.70	9.3e-64	2299	1	AK054649 Homo sapiens CDNA
gb_in:RNRG	- 993.00	1166.17	1.3e-56	3227	1	X68101 R.norvegicus trg mRN
gb_hlg:AC020073	- 953.00	1082.75	5.9e-52	66964	1	AC020073 Drosophila melano
gb_in:BC009134	+ 920.00	1079.55	9.0e-52	2998	1	BC009134 Mus musculus, Stim
gb_pat:BC015018	+ 901.00	1059.30	1.2e-50	2413	1	BC015018 Homo sapiens, Clon

seq_name: gb_pat:AX172874

seq_documentation_block:

LOCUS	AX172874	6828 bp	DNA	linear	PAT 03-JUN-2001
DEFINITION	Sequence 1 from Patent WO0142297.				
ACCESSION	AX172874				
VERSION	AX172874.1				GI:14597911
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
TITLE	Lu, P., Garman, J.D. and Candia, A.F.				
JOURNAL	Clasp-3 transmembrane protein				
FEATURES	Patent: WO 0142297-A 1 14-JUN-2001;				
source	Arbor Vita Corporation (US)				
	Location/Qualifiers				
	1..6828				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/note="full-length human CLASP-3 CDNA"				
	22..6294				
	/note="human CLASP-3"				
	/codon_start=1				
	/protein_id="CAC43726.1"				
	/db_xref="GI:14597912"				
	/translation="MAERRAPROKISRTVAAYFKVKGISGSGSPOLKLNITNGIS				
	HTTPTLEADVVDLELYLTHPLAVDSGLPRLDIEPPDIEVYSPRCRTLVSA				
	YVESEMDPHVNDCLRSYTEMDVAIVIRYHNLGTFNNTLDKQERQGLKQVFS				
	DEAPGNSYDODDGLKRRMSIDTPPGSNAGSIFDKNLSPALLPMLDPRNES				
	IDRONDKSRHKEFLPALSPEEPIERLSPVDPKHFQGLLKLCTSKETI				
	ETLEPFASDIAVDKKEKISNPFEDNSOMGLAPHPAPATITLARSALFITY				
	YTSQVFLYIKELKVLQGDGECEAPMTIKEDADYNNKLEKLSQAPQORIG				
	KTRPFWAMTALHMLNIVSSAGLSLSDSTVELEISTEKGSSERNSSIVGRSLERT				
	TSQDACLNTSPFATLVYTFNFKQDGLDDEDELYKFLAMRRSSVLRIRLTQ				
	LKIDISPAENHFCITPELLOVYKLPDSRAPTELEIFPARVYVNTYRMLTY				
	YPOSINPENNROSARNTIVKVOEMGEPSNAMPVIFEGSSCSFKAYVAAYVHN				
	SPDHEELKYKLPATLTDHNLHLPFYVSOQKONPLFPVCTYTPMLQNRKLT				
	GOFCPLVSLERPPQATVSLSPVPLPGKKVNDHNGVNVAVSSITHDDPYLDR				
	FALVLADEHLEPVARIGDMRLMENLENLSSISALNSQLEPVRLHLIDLIL				
	LVTRPVLAGQVNLGOASFEMASITIRLKHNLGNDDQGRNSLASLTYHYFRIP				
	NIYPPSSPGFGIGSVHYATMARSAVRASLINRSLNSNSDTSQPTSPDE				
	VRSITIGSLKDSNSWVNGGPKAPMGSNPSAESQADMRSCNRSHTESNIFL				
	QTLTRLPKTLRHFHEALQWYVCGSVRESALQOAMPFELMKSMYHNHLYPMKLE				
	ARKRSRFRERMDITLALVSTIASDVSFQDIDBVRLLSLAFNLDSLWDKQ				
	FEVSLIKSCYKQVSSKLSLPSPVSLRLDIFKILCSHHVYVLLPCLSLPPAS				
	PFSVSATSQSGFSTWQODIKNMLSELSPFQOQYLAGVLTBLAVLIDDAEG				
	LEGRLKRYLNMVNLHSSIDSPRSDQIARQVAMVLPIGILMETPQLYPTET				
	HNORRPTCIATDDYSESGMSISQTVAMLAGVSPVLTDRGSLFTSTGRQHTP				
	SNASSRSLITCLMLKNADEVLOKFTDLSYQNLNLDLILYICVCEPKKKKVF				
	EKNNSITRKSMDKRAKLEBALIGSITGRQEVNRSRQGLRSRBSGAFSGQENRMR				
	KQMTWROKTEMLDKSRAELIEHALIDENLEALILDLLELVQVAFVSEKSI				
	LEGVLKVLHSAACNOSAVYLOHCFEATRALVSKPELLFEETEOCDNLRLRHC				
	SSSITIRSHPSASLYLIMROKFEIGNNFAVARKQVPSLSIVQTSQNFEEFLRS				
	IKTITLVEAEDELRETFPPDOVLVNLNHLISDPTVYKMEHODPMLDLIMVRIA				
	KQYTSQPREPITLQWMAKHSRSHNARACIINHSAVAVKVEYSMLERKXVLYGVC				
	PTONISSVNLERSAVSDVSPDEGICSGKFTVSGVLGLBQAAFSMAQNEAV				
	NEVYVLLPITHANRDAKLSLTISKLEDAFSLIHQSTGEMRIVKFRGFTGRT				
	GLDQDEEVYKPAATTKLAEISHRLEGVGEFGVADVAKSDPVKCDLQNKAY				
	IQITVYPERYFDYEMKDRITLYDDKYNLNRREYCTPTFLDGRHLEDEQKRTTIL				
	THAPRPIKTRVNVNKHKEIILTPLEVLEDMOKTOELARATHODPRLMVLIO				
	GVGTVTNGRLEVAQVQLSEIPSDPKLFRHNHNRKLRLOKFTKRCEDALKKNSLIG				
	PQKRYQRLKGLTSP"				

BASE COUNT 2082 a 1384 c 1432 g 1930 t

ORIGIN

alignment_scores: Quality: 10815.00 Length: 2990
 Ratio: 5.175 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-737-246-2 x AX172874 ..

Align seg 1/1 to: AX172874 from: 1 to: 6828

```

1 MetAlaGluArgAlaPheAlaGlnIleSerArgThrValAlaI 17
  |||
22 ATGGCCGAGCGCCGCCCTTCCGCCAGAAAGATCAGAGAACGGTGCAGC 71
  |||
17 aGluValArgLysGlnIleSerGlyGlnTyrSerGlySerProGlnLeuL 34
  |||
72 CGAAGTTAGGAAGCAGATCTCCGGACAAATATAGTGTCTCCCAACTGC 121
  |||
34 euLysAsnLeuAsnIleValIleGlyAsnIleSerHisIsthrThrValPro 50
  |||
122 TCAAAACCTTAATATTTGGCAATATATCCCATCACACACAGGTGCC 171
  |||
51 LeuThrGluAlaValAspProValAspLeuGluAspTyrLeuIleThrHi 67
  |||
172 CTTACCGAGCAGTAGATCCAGTGGATTGGAAAGATTACCTCATTCACA 221
  |||
67 sProLeuAlaValAspSerGlyProLeuArgAspLeuIleGluPheProp 84
  |||
222 TCCTTGGCGTGAGATTCTGGGCTTTACGGGATTGTGATGAAATTCCTC 271
  |||
84 roAspAspIleGluValValTyrSerProArgAspCysArgThrLeuVal 100
  |||
272 CAGATGATATTGAAGTTGTTTATAGTCTCGGACTGCAGACTCTTGTT 321
  |||
101 SerAlaValProGluGluSerGluMetAspProHisValArgAspCysI 117
  |||
322 TCAGCTGACTCGAAGAAAGTGAATGATCCACATGTTAGAGACTGTAT 371
  |||
117 eArgSerTyrThrGluAspThrAlaIleValIleArgLysTyrHisLysL 134
  |||
372 AAGAAGTTTATACAGAAAGACTGGGCAATTGTCAAGAAAATATCATTAAT 421
  |||
134 euGluTyrGlyPheAsnProAsnThrLeuAspLysGluLysGluArgGln 150
  |||
422 TGGGAACAGAGATTATATCCCATATACATTAGATAACACAGAAAGCGCA 471
  |||
151 LysGlyLeuProLysGlnValPheGluSerAspGluAlaProAspGlyAs 167
  |||
472 AAAGGTTTGCAAACAAAGTTTGAATCTGATGAAGCTCCGATGGCAA 521
  |||
167 nSerTyrGlnAspAspGlnAspAspLeuLysArgArgSerMetSerIleA 184
  |||
522 CAGCTACCGAGATGATCAAGATGACCTTAAAGAGCTTCAATGTCAAAG 571
  |||
184 sPAspThrProArgGlySerTyrAlaCysSerIlePheAspLeuLysAsn 200
  |||
572 ATGATATCCCAAGGGGTAGCTGGCGCTGTAGTATCTTGACTTGAAAAAT 621
  |||
201 SerLeuProAspAlaLeuLeuProAsnLeuLeuAspArgThrProAsnG 217
  |||
622 TCACCTCCGATGCTTGGCTTCCCAATTACTTGTATCGAACTCCAAATGA 671
  |||
217 uGluIleAspArgGlnAsnAspAspGlnArgLysSerAsnArgHisLysG 234
  |||
672 AGAAATAGACCGTCAGAAATGATGACCAAGAAATCAAAACGTCACAAG 721
  |||
234 IuLeuPheAlaLeuHisProSerProAspGluGluProIleGluArg 250
  |||
722 AACCTTTTCTTGTCATCATCACAGATGAGGAAGAACCAATATGAACGG 771
  |||
251 IeuSerValProAspIleProLysGluHisPheGlyGlnArgLeuLeuVal 267
  |||

```

```

772 CTATAGTGTCTGATATACCCAAAGAACATTTTGGTCAAGACTTCTGT 821
  |||
267 IlycCysLeuSerLeuLysPheGluIleGluIleGluProIlePheAla 284
  |||
822 AAAATGCTTATACACTCAAGATTGAAATTTGAATTTGAACCATTTTGGCAA 871
  |||
284 eLleuAlaLeuTyrAspValLysGluLysLysLysIleSerGluAsnPhe 300
  |||
872 GTTGGCTTATATGATGTCAGGAAAGAAAGAAAGATTTCAGAAACTTT 921
  |||
301 TyrPheAspLeuAsnSerGluGlnMetLysGlyLeuLeuArgProHisVal 317
  |||
922 TATTTTGACCTTAATTTCTGAGAGATGAAGGGTGTGTACGTCCACATGT 971
  |||
317 IProProAlaAlaIleThrThrLeuAlaArgSerAlaIlePheSerIleT 334
  |||
972 ACCACCTGTGCGCATTTACTACCCCTGGCAAGATACGCAATTTTTCATCA 1021
  |||
334 hTyrProSerGlnAspValPheLeuValIleLysLeuGluLysValLeu 350
  |||
1022 CTATTCTTCCCAAGATGTTTCTTCTGTATTAAGCTAGAAAAGTCTTA 1071
  |||
351 GlnGlnLysAspIleGlyGluCysAlaGluProTyrMetIlePheLysG 367
  |||
1072 CAGCAAGGAGACATTTGAGAGGTGGCAGAACCATATATGATTTTCAAGA 1121
  |||
367 uAlaAspAlaThrLysAsnLysGluLysLeuGluLysLeuLysSerGln 384
  |||
1122 AGCAGATGCGCACCAAGAAATGAAGAAAACCTGAGAAAGTCAAG 1171
  |||
384 IAspAspGlnPheCysGlnArgLeuGlyLysTyrArgMetProPheAlaTrp 400
  |||
1172 CAGATCGATTTTGGCAAGACTTGGGAAATATCGCATCTTTTGCTGG 1221
  |||
401 ThrAlaIleHisLeuMetAsnIleValSerSerAlaGlySerLeuGluArg 417
  |||
1222 ACTGCATTCATTTAATGAATATTTGTAGCAGTGTGGGAGTTTGGAAAG 1271
  |||
417 gAspSerThrGluValGluIleSerThrGlyGluArgLysGlySerTrps 434
  |||
1272 AGATTCTACAGAAATGAAATCAGTACTGGAGAACGAAAGGGCTCTTGGT 1321
  |||
434 eArgLysArgAsnSerSerIleValGlyArgArgSerLeuGluArgThr 450
  |||
1322 CAGAGAGAGAGAAATTTAGTATTTGTTGGCAGACGATCACTTGAAGGACA 1371
  |||
451 ThrSerGlyAspAspAlaCysAsnLeuThrSerPheArgProAlaThrLe 467
  |||
1372 ACAAGTGGAGATGATGCTGTGAACCTTGACGAGCTTTCGACCACCTACTGT 1421
  |||
467 uThrValThrAsnPhePheLysGlnGluGlyAspArgLeuSerAspGlu 484
  |||
1422 CACAGTGACAAATTTTTTTTAAGCAGAGAGAACCCCTTAAAGTAGAAGA 1471
  |||
484 sPLeuTyrLysPheLeuAlaAspMetArgArgProSerSerValLeuArg 500
  |||
1472 ATCTCTCAAAATTCCTTGCTGATATGAGAAAGCCATCTTGTGCTTACGG 1521
  |||
501 ArgLeuArgProIleThrAlaGlnLeuLysIleAspIleSerProAlaArg 517
  |||
1522 CGACTAGACCTAATTTACAGCTCAGCTCAAGATAGACATTTCTCCCGCAC 1571
  |||
517 oGluAsnProHisTyrCysLeuThrProGluLeuLeuGlnValLysLeu 534
  |||
1572 TGAATAATCCCATTAATTTGCTTAACCTCGGAGCTGCTTCAAGTAAACCTTT 1621
  |||
534 YrProAspSerArgValArgProThrArgGluIleLeuGluPheProAla 550
  |||
1622 ACCCTGACAGTAGAGTTAGCTTACCAAGAGAAATCTTAGAGTTTCCCGCA 1671
  |||
551 ArgAspValTyrValProAsnThrThrTyrArgAsnLeuLeuTyrIleTyr 567
  |||
1672 AGGATGTTTATGTTCAAACACTACTTACAGAAATCTTCTTACATATA 1721
  |||

```

567 rProGlnSerLeuAsnPheAlaAsnArgGlnGlySerAlaArgAsnIleT 584
|||||
1722 CCCTCAGAGTCTTAATTTGGCAATCGCAAGTCTCTGTAAGAAATATATA 1771
584 hrValLysValGlnPheMetTyrGlyLysAspProSerAsnAlaMetPro 600
|||||
1772 CAGTGAAGTCCAGTTTATGTAAGAGAGATCCAGCAATGCCATGCCG 1821
601 ValIlePheGlyLysSerSerCysSerGlnPheSerLysGlnAlaTyrTrh 617
|||||
1822 GTAATCTTTGGTAATCTAGCTGTTCAGAAATTTCAAGGAAGCTATAC 1871
617 rAlaValValTyrHisAsnArgSerProAspPheHisGlnGluIleLysV 634
|||||
1872 ACCCGTAGATATCATTAACAGGTCCTCGATTTTCATGAGAAATCAAG 1921
634 alLysLeuProAlaThrLeuThrAspHisHisLeuLeuPheThrHe 650
|||||
1922 TTAAAGCTTCCTCTACTTTAACTGAACATCAACACTTCTTTTAACTTTT 1971
651 TyrHisValSerCysGlnGlnLysGlnAsnThrProLeuGluThrProVa 667
|||||
1972 TATCATGTTAGTTGTTCACAAAAACAAATACTCCTCTTGAACACCACT 2021
667 lGlyTyrThrTrpIleProMetLeuGlnAsnGlyArgLeuLysTrhGlyG 684
|||||
2022 TGGATATACATGATACCAATGCTTCAGAAATGACGGTTGAAGACTGGCC 2071
684 lPheCysLeuProValSerLeuGluLysProProGlnAlaTyrSerVal 700
|||||
2072 AGTTTGGCTTGCAGCTCATTTGGAAGAAACCAACAGGCTTATCTGTGA 2121
701 LeuSerProGluValProLeuProGlyMetLysTrpValAspAsnHisLys 717
|||||
2122 CTGTCTCCGAGAGTCTCTACTGCGCATGGAATGGTAGATAAATACAA 2171
717 sGlyValPheAsnValGluValAlaValSerSerIleHisTrhGlnA 734
|||||
2172 AGGTGTTTAAATGTTGAAGTGTGTGCTGTTGCTATCCATACACAAG 2221
734 sPProTyrLeuAspLysPhePheAlaLeuValAsnAlaLeuAspLunHis 750
|||||
2222 ATCTTATCTTGACAAATTTTGGCTGTGTCATGCAATGCTGTGATTAAC 2271
751 LeuPheProValArgIleGlyAspMetArgIleMetGluAsnAsnLeuG 767
|||||
2272 CTGTTCGCCAGTCCGAATTTGGGACATGCGAATCATGAAAAATTAACTTAGA 2321
767 uAsnGluLeuLysSerSerIleSerAlaLeuAsnSerSerGlnLeuGluP 784
|||||
2322 AAATGAAATGGAAGAGAGATTTTCAAGCACTGAATTCATCCAGCTGGAGAC 2371
784 roValValArgPheLeuHisLeuLeuLeuAspLysLeuIleLeuLeuVal 800
|||||
2372 CAGTGGTCCGATTTCTCATCTTTCGTAGATTAACGTATTACTTTTAACT 2421
801 lLeuArgProProValIleAlaGlyGlnIleValAsnLeuGlyGlnAlaSe 817
|||||
2422 ATTAGACCTCTGTCTACTGTGCCAAATAGTTAACTTAGTCAAGCATC 2471
817 rPheGluAlaMetAlaSerIleIleAsnArgLeuHisLysAsnLeuGluG 834
|||||
2472 TTTTGAAGCATGGCATCAATATTAATGACTTCACAAAAAACTGGAGAG 2521
834 lAsnHisAspGlnHisGlyArgAsnSerLeuLeuAlaSerTyrIleHis 850
|||||
2522 GAAATCATATGACAGATGGCAGAAACAGCCTTCTTGATCATATATATCAT 2571
851 TyrValPheArgLeuProAsnThrTyrProAsnSerSerSerProGlyPyr 867
|||||
2572 TATGTTTCCGCTACCAAAATACTTACCTTAATCATCATCATCACAGGTCC 2621

867 oGlyGlyLeuGlyGlySerValHisTyrAlaThrMetAlaArgSerAlaY 884
|||||
2622 TGGGGTTTTGGAGAGATCACTGATTAATGCCAATAGGCTAGATCTGCG 2671
884 alArgProAlaSerLeuAsnLeuAsnArgSerArgSerLeuSerSerAsnSer 900
|||||
2672 TGAGACCTGCCAGCCTTAATTTAAATCGTTTCTCGAAGCCTTAGTAATAGC 2721
901 AsnProAspIleSerGlyThrProThrSerProAspAspGluValArgSe 917
|||||
2722 AATCAGATATATCTGGAGATCCACAGCTCACCAGATGATGAATCTGCATC 2771
917 rIleIleGlySerLysGlyLeuAspArgSerAsnSerTrpValAsnTrhG 934
|||||
2772 AATCATCGGAGATTAAGGTTTAACTCGCTCAATTCCTGGGTAAACACTG 2821
934 lGlyProLysAlaAlaProTrpGlySerAsnProSerProSerAlaGlu 950
|||||
2822 GTGGTCCAAAAGCTGCCCATGGGAGTCCAAACCCAGTCCAGTGCAGAA 2871
951 SerThrGlnAlaMetAspArgSerCysAsnArgMetSerSerHisTrhG 967
|||||
2872 TCACACAGAGCTATGATTCGAAAGTTGTAATCGTATGCTTCGACACACGA 2921
967 uThrSerSerPheLeuGlnThrLeuThrGlyArgLeuProThrLysLysL 984
|||||
2922 GACGTCAAGTTTCTTACAAACATTAACGGGAGCGCTTACCAACTAAAAAGC 2971
984 eupHeHisGlnGluLeuAlaLeuGlnThrValValCysSerGlySerVal 1000
|||||
2972 TTTTTCACAGAGAGCTGGCTTTCAGTGGGTGTGTTCACAGTGGCAGCGT 3021
1001 ArgGlnSerAlaLeuGlnGlnAlaTrpPhePhePheGluLeuMetValLys 1017
|||||
3022 CGGGAATCAGGCTTTCGACAAAGCCTGCTCTTTTTCATTAATTAATGGA 3071
1017 sSerMetValHisHisLeuTyrPheAsnAspLysLeuGluAlaProArgL 1034
|||||
3072 GAGCATGTGTCACCATTTATATGATTAACCTTGAAGGCTCCAAAGGA 3121
1034 ySerArgPheProGluArgPheMetAspAspIleAlaAlaLeuValSer 1050
|||||
3122 AAAGTCTGTTTCCAGAACGTTTCATGGATGACATTCGAGCTCTTGTCAGC 3171
1051 ThrIleAlaSerAspIleValSerArgPheGlnLysAspThrGlnMetVal 1067
|||||
3172 ACGATGTCTAGTGAATAGTTTCACGATTTTCAGAGAGACACAGAAATAGGT 3221
1067 lGluArgLeuAsnThrSerLeuAlaPhePheLeuAsnAspLeuLeuSerV 1084
|||||
3222 TGAGAGACTCAATACAGCCTTGATGCTTTCCTCAATGATCTGTGCTG 3271
1084 alMetAspArgGlyPheValPheSerLeuIleLysSerCysTyrLysGln 1100
|||||
3272 TTATGGACAGAGAGATTTGTTTAACTTAAAGTCTCTTAAACAG 3321
1101 ValSerSerLysLeuTyrSerLeuProAsnProSerValLeuValSerLe 1117
|||||
3322 GTGTCTTCAAAAGCTTACTATTCATCCGAATCCAGTGTCTGTGGTGTCTT 3371
1117 uArgLeuAspPheLeuArgIleIleCysSerHisGlnHisTyrValThrL 1134
|||||
3372 GAGGCTGATTTTCTACGAATCATCTGCAATCATGAGACACTATGTTACAT 3421
1134 euAsnLeuProCysSerLeuLeuThrProProAlaSerProSerProSer 1150
|||||
3422 TAAACTTAACCTCGAGCTTACTTACTCACCTGCATCTCCATCACTCTCT 3471
1151 ValSerSerAlaThrSerGlnSerSerGlyPheSerThrAsnValGlnAs 1167
|||||
3472 GTTCTCTTCGACACATCTCAGAGTGTGGAATTTTCTTCGATGTACAGAGA 3521
1167 pGlnLysIleAlaAsnMetPheGluLeuSerValProPheArgGlnGlnH 1184

```

|||||
3522 CCAAAAGATTGCAATATGTTGAATATCGCGCTTCGCCCAACAGC 3571
1184 lstrleuValaglyLeuValleuthrGluLeuAlaVallleuAspPro 1200
|||||
3572 ATATATTGGCAGACCTTGTTAAACAGACCTGGCTGTCATTTTAAACCT 3621
1201 AspAlaGluGlyLeuPheGlyLeuHisLysLysVallleAsnMetValHl 1217
|||||
3622 GATGCTGAAGAGACTGTTGGATTGCTAAGAAAGTCATCATATATGTTACA 3671
1217 sasnleuLeuSerSerHisAspSerAspProArgTyrSerAspProGlnI 1234
|||||
3672 CAATTACTCTCCAGTCACGACACTAGACCCGGGTACTGACCTCAGACA 3721
1234 leuValAlaArgValAlaMetleuTyrLeuProleuIleGlyIleIleMet 1250
|||||
3722 TAAAGGCTGAGTGGCCATGTTGATCTACCTGATGTTGATATATCATG 3771
1251 GluThrValProGlnleuTyrAspPheThrGluThrHisAsnGlnArgG1 1267
|||||
3772 GAAACCTGACCTGACGTGATGATTTTACAGAAACTCACATCAACAGAG 3821
1267 yArgProIleCysIleAlaThrAspAspTyrGluSerGluSerGlySerM 1284
3822 AAGACCAATTTGATAGCCACATGATGATATGAAAGTGAGAGCGGAAGTA 3871
1284 etIleSerGlnThrValAlaMetAlaIleAlaGlyThrSerValProGln 1300
|||||
3872 TGAATAGCCAGACCGTTGCCATGGCAATGCGAGGACATCGGTCCCTCAA 3921
1301 leuThrArgProGlySerPheleuLeuThrSerThrSerGlyArgGlnHl 1317
|||||
3922 CTAAACAAGGCTGCGAGTTTCTCTCCACAGTCACAGAGGCGACGACACA 3971
1317 strThrPheSerAlaGluSerSerArgSerLeuLeuIleCysleuLeuT 1334
|||||
3972 CACTACCTTTTCAGCAGAAATCAAGTCGAGAGCCTTTTGATCTGTCTACTT 4021
1334 rValleuLeuLysAsnAlaAspGluThrValleuGlnLysThrPheThrAsp 1350
4022 GGGTCTCTAAAATGCAATGAAACAGTCTTACAGAAAGTGTTTACAGAT 4071
1351 leuSerValleuGlnleuAsnArgleuLeuAspLeuLeuTyrLeuCysVa 1367
|||||
4072 CTCTCAGTCTTGACAGTAAACCGGCTATTAGATCGCTTTATCTCTGCT 4121
1367 lSerCysPheGlyTyrLysGlyLysLysValPheGluArgMetAsnSerL 1384
|||||
4122 GTCTTGCTTTGAGTATAAAGGAAAAAGTCTTGAACGAATGATACCT 4171
1384 euThrPheLysLysSerLysAspMetArgAlaLysleuGlnGlnAlaIle 1400
4172 TGAACCTTTAAGAAATCAAAAAGCATGAGAGCAAAAGCTTGAAGAACCTAT 4221
1401 leuGlySerIleGlyAlaArgGlnGlnMetValaArgArgSerArgG1 1417
|||||
4222 CTTGGGAGCATAGTCCAGCGCAAGAAATGTTACGCGCAAGCGCGAGACA 4271
1417 nleuGlnArgSerProSerGlySerAlaPheGlySerGlnGlnuAsnLeuA 1434
4272 GCTCGAGAAAGCCCATCTGGAAGTGCTTGGAGTCAAGAAAAATTTGA 4321
1434 rGTTPArgLysAspMetThrHisTPArgGlnAsnThrGluLysleuAsp 1450
|||||
4322 GGTGGGGAAGATATGACTCTACGCGTCAAAACACAGAGAAAGCTTGAC 4371
1451 LysSerArgAlaGluIleGlnHisGlnAlaLeuIleAspGlyAsnLeuAl 1467
|||||
4372 AAATCAAGAGCAGATTTCAACACAGAGCACTGATTGATGGAACCTGCG 4421
1467 aThrGlnAlaAsnLeuIleIleLeuAspThrLeuGluIleValValGlnT 1484
|||||

4422 TACAGAGCAAAACCTAAATTCATTTTACATACATTAGACATTGTTGTCAGA 4471
1484 hrValSerValThrGluSerLysGluSerIleleuGlyGlyValleuLys 1500
|||||
4472 CCGTTTCTGTAAACGAATCCAAAGAGAGCATTTCTGGTGAGTGCCTAAA 4521
1501 ValleuLeuHisSerMetAlaCysAsnGlnSerLysValaTyrleuGlnHl 1517
|||||
4522 GTGCTACTACACAGATGGCTGTACCAAAAGTGCAGTTTATCTACACACA 4571
1517 sCysPheAlaThrGlnArgAlaLeuValSerLysPheProGlnleuLeuP 1534
|||||
4572 CTGTTTGTCTACACAGAGAGCCTTGTTCAAGTTTCTCGAAGCTGTAT 4621
1534 heGluGluGluThrGluGlnCysAlaAspLeuCysleuAlaGlyleuLeuArg 1550
4622 TTGAAGAGAGACAGACAGTGTGCTGATTATGCTTCAGGCTTCTCGA 4671
1551 HisCysSerSerSerIleGlyThrIleArgSerHisProSerAlaSerIle 1567
4672 CACTGTACAGTACCATCGGTACATATCGGTACACCCCGAGTGCCTGCT 4721
1567 uTyrleuLeuMetArgGlnAsnPheGluIleGlyAsnAsnPheAlaArgV 1584
|||||
4722 TTACCTACTAATGAGCCAAACCTTGAGATTGGGAATTAACCTTGCCAGG 4771
1584 aLysMetGlnValProMetSerLeuSerLeuValIleGlyThrSerGln 1600
4772 TTTAAATGCAAGTACCAATGCACTATCCCTTGTTGGGCGACATCTCAG 4821
1601 AsnPheAsnGluGluPheLeuArgArgSerLeuLysThrIleLeuThrTy 1617
4822 AATTTTATGAGAATCTTAAGACGTTCTCTAAAGCTATATTGACATA 4871
1617 rAlaGluGlnAspLeuGlnleuArgGluThrThrPheProAspGlnValG 1634
|||||
4872 TGCTGAAGAAGATCTGGAATTTGAAGGAACAACATTTCTGATCAGAGTCC 4921
1634 lAspLeuValPheAsnLeuHisMetIleLeuSerAspThrValLysMet 1650
4922 AGGATGTGGTTTCAATCTCATATGATTTCTTCTGATCTGGAANAAG 4971
1651 LysGlnHisGlnGlnuAspProGluMetleuIleAspLeuMetTyrArgI 1667
4972 AAGAACACACAGAGAGATCTGGAATTTGATTTGATGATTAATGACAGAA 5021
1667 eAlaLysGlyTyrGlnThrSerProGluArgleuThrTrrPleuGlnAsnM 1684
5022 TGCCAAAGGTTTACAGACCTCTCCAGAGCATTTGACCTGGTTCAGACACA 5071
1684 etAlaGlyLysHisSerGluArgSerAsnHisAlaGlnAlaGlnCys 1700
5072 TGGCAGGCCAAGCACTCGAAGCAAGCAATCATGCTGGAAGCTGCAGAGT 5121
1701 leuValHisSerAlaAlaLeuValAlaGluTyrLeuSerMetleuGlnuAs 1717
5122 CTATCTCACCTACAGACACATTTGCTGCAATATTGAGCATGCTGGAGACA 5171
1717 PArgLysTyrleuProValGlyCysValThrPheGlnAsnIleSerSerA 1734
5172 CCGGAATATATCTTCCCTGGGAGTGTGAACATTTCAAGAAATTTCAATCA 5221
1734 snValleuGluGluSerAlaValSerAspAspValValSerProAspGlu 1750
5222 ATGTTTGAAGAAATGCGGCTCCAGATGATGCTGATCTCCAGATGAAGA 5271
1751 GluGlyIleCysSerGlyLysTyrPheThrGluSerGlyLeuValGlyLe 1767
5272 GAAGGTATCTGCTTGAAAAATATCTTACTGAGTCAAGAGACTTGTGGAGAT 5321
1767 uLeuGluGlnAlaAlaAlaSerPheSerMetAlaGlyMetTyrGlnuAlav 1784
5322 ACTGGAACAAGCAGCTGCTCTTCTCTATAGCTGGCATGTATGAAGCAG 5371

```

1784 aAsnGluValTyrLysValLeuIleProIleHisGluAlaAsnArgAsp 1800
|||||
5372 TTAATGAAGTATTACAAGTACTATTCTATTCATGACGAAGCTAATCGGAT 5421
1801 AAlaLysLysLeuSerThrIleHisGlyLysLeuGlnGluAlaPheSerLys 1817
|||||
5422 GCAGAAAGAACTATCCACAATTCATGTGTAACACTTCAAGAAAGCATTCAGCA 5471
1817 sIleValHisGlnSerThrGlyTrpGluArgMetPheGlyThrTyrPheA 1834
|||||
5472 AATGTGTCACAGACTACGCTGGAGGAGGATGTTGGCACTTATTTTC 5521
1834 rGValGlyPheTyrGlyThrLysPheGlyAspLysLeuAspGluGlnLuphe 1850
|||||
5522 GTGTGGTTTATGGAACCAAGTTCGGGGATTGGATGACAAACAATTT 5571
1851 ValTyrLysGluProAlaIleThrLysLeuAlaGluIleSerHisArgLe 1867
|||||
5572 GTTTACAAAGAGCTCGCATATAACCAACTTGCAGAGATATCTCACAGATT 5621
1867 uGluGlyPheTyrGlyLysLysPheGlyGluAspValAlaGluValIleL 1884
|||||
5622 GAGGGAATTTACGGAGAGAGATTGGAGAGATGTTGGTGAAGTAATCA 5671
1884 yAspSerAsnProValAspLysCysLysLeuAspProAsnLysAlaTyr 1900
|||||
5672 AAGACTCTAATCTGTAGACAAAGTGAATATGATCTTAACAGACATAT 5721
1901 lIleGlnIleThrTyrValGluProGlyPheAspThrTyrGluMetLysAs 1917
|||||
5722 ATTCACATTAATCTATGTGAGGACCACTTGCACATATGACATGAAGA 5771
1917 pArgIleThrTyrPheAspLysAsnTyrAsnLeuArgArgPheMetLysC 1934
|||||
5772 CAGAATCACCCTATTTCGACAAAATTTACATCTTGTGCAATTCAGTACT 5821
1934 yThrProPheThrLeuAspGlyArgAlaHisGlyLysLeuHisGluGln 1950
|||||
5822 GACACACCTTACTTACTTATGATGCGCGCCATGGGGAACCTTATGACAA 5871
1951 PheLysArgLysThrIleLeuThrThrSerHisAlaPheProTyrIleLys 1967
|||||
5872 TTCAAAAGGAAGACCAATCTGATCGTCTCATGCTTCCCTTATATATA 5921
1967 sThrArgValAsnValThrHisLysGluGluIleLeuThrProIleG 1984
|||||
5922 AACAGGGTCAATGTCATCATTAAGAGAGATCATCTTAACACCAATGG 5971
1984 LuValAlaIleGluAspMetGlnLysLysThrGlnGluLeuAlaPheAla 2000
|||||
5972 AAGTTCCTATTGAGCATGACAGAAAAAGACAGAGAGTGGCATTTGCA 6021
2001 ThrHisGlnAspProAlaAspProLysMetLeuGlnMetValLeuGlnI 2017
|||||
6022 ACACATCAGATCCCGCAGACCCCAAAATGCTTCAGATGCTACTCCAGGC 6071
2017 ySerValGlyThrThrValAsnGlnGlyProLeuGluValAlaGlnValP 2034
|||||
6072 ATCTGTAGGCACACAGTCAATCAGGGGCTTTGGAAAGTGGCCCGAGTTT 6121
2034 heLeuSerGluIleProSerAspProLysLeuPheArgHisHisAsnLys 2050
|||||
6122 TTCTGCTGAATAATACCTAGTACCCAAAGCTCTTCAGACATCATATATA 6171
2051 LeuArgLeuCysPheLysAspPheThrLysArgCysGluAspAlaLeuAr 2067
|||||
6172 CTCGCACTCTGCTTAAAGATTTTACTAAAGGTGGAAGAGCTTAAAG 6221
2067 gLysAsnLysSerLeuIleGlyProValGlnLysGluTyrGlnArgGluL 2084
|||||
6222 AAAAAATTAAGACTTAATTTGGCGGCTCAAAAAGAGATATCAAGAGGAA 6271

2084 euGlyLysLeuSerSerPro 2090
|||||
6272 TGGGAAACTATCTCGCCT 6291
seq_name: gb_pat:AX173175
seq_documentation_block:
LOCUS AX173175 6372 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 1 from Patent WO0142295.
ACCESSION AX173175
VERSION AX173175.1 GI:14598036
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 6372)
AUTHORS Lu, P., Garman, J.D. and Candia, A.F.
TITLE Clasp-7 transmembrane protein
JOURNAL Patent: WO 0142295-A 1 14-JUN-2001;
Arbor Vita Corporation (US)
FEATURES
source location/Qualifiers
1..6372
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="full length human CLASP-7 cDNA"
13..6156
/note="human CLASP-7"
/codon_start=1
/protein_id="CAC43731.1"
/db_xref="GI:14598037"
/translation="MAASERAFAHKINRTVAAEVRKQVSRERGSPPHSRCCSSILG
VPLTEVEPLDEEDVILSRPDAEPGLDVEFPADDELLOPRECTREPGIPKD
EKLDQVRAVEMYIEDMWIVRRYQVLSAAYPTPTQERQKGLAEPQVEEDASG
DERSGPDSNDRSGSPEDPRSSGASIEDLNLAADLSILSLRAEDVDNR
NETLKRQHPALILTPAPDEDEAVERCSRPREPRHFGRIILKCLSLKEFLEIEP
DLEIVIKLELQGDISECEPYNVLEKVTAKKKELEKRLAAEFCTRLGRYM
PRAMVAHLANTVSSAGOLDSDSEGERRRAPRRRRQVQDRASGDDCSGFR
PATLVNTFNRKEARLSDELEKFLADMRRESILRLRLVDTOLKIDISAPENPH
FCLSPELHIKYPDPGRGPTKEILEPPAREVYANHSYRNLLTYVPHSLMPSRGS
VNLAVRQVMTGEDPSQALPYIFGKSSCSETRAPTPVYVYHNSPEFTEERKULHP
ACVTENHLLFTFVHVSCOPRGTALETPEVGTWIPLDQHRLETRGPCLPVSVDOP
PESVLTEDVALPGRWVDGKGVSELTAVSSVHPDQPYLDLDEFTLVHLEGAEP
FLTKDPLVSEGNVEQLRSALALASPEPLASVHHLVLDKVLVLRPIISGOIV
NIGRAEFEMAHVSLVHRSLEAQAADAGHQPQAAVYVAFRLPGTEPSLPDAPV
TYQAATFLANGSRPASLYLARSKSISSNPDLAVAPGVDEVSIIILAKLHLEAL
QNVVSSAVREALIDHAMFFQLMWKSMALHLLDLSQDRIDTRKLPFGERTLDDITALY
GSGLEVEIRVHKDELAEHLNASTLAFSLDLSLDVDRGFVSLVRAHYKQVARTLOS
SNPVALILTRMEFTRILCSHEHYVTLMPCCLSPSPASPSVSTSSQSTFSQA
PPPKYTMSELSPRROOHFLAGLITLDELALAEAGAEFLHKKAISAVHSLCGH
DDIPRYAEATVAVAFELIPLISTARTPLPHEFAFGPGRASRLAASDDEGEG
DIAGTINPSVMAAINGPLAPSRASISIQGPPTARACALASBSRTLLACVIMWLK
NTEPALQWATDITLTPQGLRDLDTLCLAEFTYKKAERINSLTFKSLDKAR
LEBALIGTGAQOEVRSRERSRPGNEBNWRKSVTHMQTSORVDTKDEMEHGA
LVEGNATEASLVVDLTLEIIVQTVMLSEARSVYGAVLKVLVSLGASQALPFOHG
LATORALVSKPELLEDEEDLCADLCRLRHCSRLISTIKRTASASLYLMRQNE
IGHNRAVYMQVTMSSLSLVGTQNSSEHLRRSKTISTIAEEDMGLSDTFPAVOV
DIMFNHMTITDTYKMKHEQDEPMLIDLMRLIARGYSGSPRTTWNAGKHAEL
GHAABEQCMVHAALVAYEYLALLEDQRLHPVGCSPONISSNVTEESATSDILSPD
EEGFCSGKHFTELGLVLEQNAAGFTYMGVLYEAVNEVYKNTLPLEHRYKRIAAV
HOKLOEAPFKIHOSGWERVFGTYFRGFGAHHGDDLEDEEVYKESYLKLEIISH
RLEEFYTERFGDVEEIIKDSYVDRSKLDSOKAYIQITYVEVPEYELKDRVYFPD
RVNGARTLEFCTPPTPDGARGHELPEOKHRRKTLSTIDHAPVYIKTRIVCHREETVLP
PVEVATIDEMOKRTRELARFTEQDDPPDAKMLQWVLOGSVGPYVNOGSLFEVAQFLAEP
EDPKLFRHNKRLRLCKPCKCEDALRNKRLISLPDOKEIYHRELERNYCRLEALOP
LITQRLPQIMATPGLRNSINRASFRADLT"

BASE COUNT 1273 a 2050 c 1849 g 1200 t
ORIGIN
alignment_scores:
Quality: 7136.50 Length: 2090

Ratio: 3.998 Gaps: 11
Percent Similarity: 85.407 Percent Identity: 65.311

alignment block:

US-09-737-246-2 x AX173175 ..

Align seg 1/1 to: AX173175 from: 1 to: 6372

```

2   AlagluararqalaphaelaglnlylleserArqThrValAlaIaagi 18
   ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
22  TCCGAGCGCGCGCGCTTCGGGCAAGATCAACAGACGGTGGCCGCACA 71
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
18  uValArglysglnIleSerGlyIntYrSerGlySerProglInleuL 35
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
72  GGTCCGGAACAGAGTGTCCCGGAGACGAGTGGCTCCCCCACTCCAGA 121
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
35  ysasnleuAnIleValGlyAsnIleSerHisThrThrValProleu 51
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
122 GG.....CGCTGCAGCAGCTCCCTGGGGGTCCCACTG 153
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
52  ThgluAlaValasProvalaspleuGluAspTyrLeuIleThrHisPr 68
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
154 ACTGAAGTTTCGACCCCTCGACTTTCGAGATGTACTTCTGAGCCGGCC 203
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
68  oleuAlaValasPserGlyProleuArgAspleuIleGluPheProPro 85
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
204 ACCAGATGCTGAGCCCGGGCCCTCAGGGACCTGTAGAAATCCCGCCTG 253
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
85  spAspIleGluValValYrSerProArgAspCysArgThrLeuValSer 101
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
254 ATGACTGGAGCTGCTGTCGACGCCCGGGAATGCGGACACGAGGCC 303
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
102 AlaValProgluSerGluMetAspProHisValArgAspCysIleAr 118
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
304 GCGATCCCGCAAGATGAAAACTGATGCCCCAGGTAGGGCCGCGGTGA 353
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
118  gSerTyrThrGluAspTyrAlaIleValIleArgLysTyrHisLysLeu 135
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
354 GATGTATATGAGACTGGGTCAATGTCCACAGAAAGTATCAGTACTGA 403
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
135  LyrThrGlyPheAsnProAsnThrLeuAspLysGlnLysGluArgGln 151
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
404 GTGCACATACAGCCCGTCAACACACAGACAGCGGGAGCGACAGAG 453
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
152  GlyLeuProLysGlnValPheGluSerAspGluAlaProAspGlyAsn 168
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
454 GGGCTCCCGCCAGGCTCTTGAAGCAGATGCTTCGGAGACGAGAGGTC 503
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
168  rTyrGlnAspAspGlnAspAspLeuLysArgArgSerMetSerIleAsp 185
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
504 CGGCCCTGAGGACTCGAATGACTCCGGGTGGCTCGGGCTCCCGGAG 553
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
185  spThrProArgGlySerTyrPalacyserTlePheAspLeuysAsnSer 201
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
554 ACACCCCTGAAAGCAGTGTGCTTCAGCATCTTCACCTGAGGAACCTG 603
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
202  LeuProAspAlaLeuLeuProAsnLeuLeuAspArgThrProAsnGlu 218
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
604 GCAGCTGACATCTGCTCCCTCTCTGCTAGACGGGGCGGCCCCAGACA 653
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
218  uIleAspArgGlnAsnAspAspGlnArgLysSerAsnArgHisLysGlu 235
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
654 TGTGGACCGCGGCAATGAACCTTCGACGCGACACGCGGCCCGGGCC 703
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
235  euPheAlaIleuHisProSerProAspGluGluGluProIleGluArg 251
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
704 TGGTCAACCTTACCCGGCACTGACGAGAGATGAAGCCGTGAAGCGCT 753
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
252  SerValProAspIleProLysGlnHisPheGlyGlnArgLeuValYr 268
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
754 AGCGCGCCAGAGACACCGCGCAGCATTGGACAAAGAGATCTTGGTCA 803
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
268  scYsIeuserLeuLysPheGluIleGluIleGluProIlePheAlaSer 285

```

```

804  GGTCTGTGCTCAAGTTCGAGATGGAATGAGCCCATCTTGGAGTCT 853
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
285  euAlaLeuTyrAspValLysGluLysLysLysIleSerGluAsnPheTyr 301
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
854  TGGCTCTGTATGATGTCCGGGAGAAAAAGAGATCTCGAGAACTGTAC 903
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
302  PheAspLeuAsnSerGluGluMetLysGlyLeuLeuArgProHisValPr 318
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
904  TTCGACTCGAACTCGGACTCCATGAGGGGCTGCTTCGGCTCATGGCAC 953
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
318  oProAlaAlaIleThrThrLeuAlaArgSerAlaIlePheSerIleThr 335
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
954  CCACCTCGCATCTCCACCTGCGCCGCTTCGCATCTCTGTGACCT 1003
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
335  YrProSerGlnAspValPheLeuValIleLysLeuGluLysValLeuGln 351
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1004  ACCCTTCACCTGACATCTTCCTGTGCATCAAGTTGAGAAAGTCTTCAG 1053
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
352  GlnGlyAspIleGlyGluCysAlaGluProThrMetIlePheLysGluAl 368
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1054  CAAGGGGACATCAGTGAAGTGTGAGCCTTACATGGTGTGAAGAAGT 1103
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
368  aAspAlaThrLysAsnLysGluLysLeuGluLysLeuLysSerGlnAla 385
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1104  GGACACACCCCAAGAAACAAGAGAGTAGAGACCTGGCGCTGGCGGCG 1153
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
385  spGlnPheCysGlnArgLeuGlyLysTyrArgMetProPheAlaTyrThr 401
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1154  AGCAGTTCTGCACCCGCTGCGCGCTACCCGATGCCCTTCGCTGAGC 1203
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
402  AlaIleHisLeuMetAsnIleValSerSerAlaCysSerLeuGluArgAs 418
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1204  GCGGCGACTTGCCACATCTGTAGCAGCGCTGGGACGCTGGACCGGA 1253
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
418  pSerThrGluValGluIleSerThrGlyLysArgLysGlySerTyrPser 435
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1254  CTCTGAC.....TCGGAGGGGAGCGCGCGGACGCTGGAGACG 1291
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
435  LuArgArgAsnSerSerIleValGlyArgArgSerLeuGluArgThrThr 451
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1292  ACCGCGCGCGT.....CGGGGCGCCGAGACCGGGCGCAGT 1326
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
452  SerGlyAspAspAlaCysAsnLeuThrSerPheArgProAlaThrLeuTh 468
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1327  AGTGGGGACGACGCGCTGAGCTTCTGTGGCTTCGCTCAGCGACGCTAAC 1376
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
468  rValThrAsnPhePheLysGlnGluLysAspArgLeuSerAspGluAsp 485
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1377  TGTCACAAACTTTCTTAAAGCAGAGGCTGAGCGCATCAGTACGACGAG 1426
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
485  euTyrLysPheLeuAlaAspMetArgArgProSerSerValLeuArgArg 501
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1427  TCTTCAAATCTCTGCTGACATGAGGGGCGCGCTGCTCCGCTCGCGGCA 1476
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
502  LeuArgProIleThrAlaGlnLeuLysIleAspIleSerProAlaProgl 518
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1477  CTAGCTCTGTGACTGCCACCTCAAGATCGACATTTCTCGGCTCTCTGA 1526
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
518  uAsnProHisTyrCysLeuThrProGluLeuLeuGlnValLysLeuYrP 535
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1527  AAATCCCACTTGTGCTCTCCCTGAGCTGCTTCATATCAAGCCCTTACC 1576
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
535  roAspSerArgValArgProThrArgGluIleLeuGluPheProAlaArg 551
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1577  CGGACCCCGAGGGCGCGCCACCAAGAGATTCGTGGAGTTCCCGCGCGC 1626
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
552  AspValTyrValProAsnThrThrTyrArgAsnLeuLeuYrIleTyrP 568
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1627  GAAGTCTATGCCCCCAATACACGCTACAGAACCTGCTGACGTGTACCC 1676
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
568  oGlnSerLeuAsnPheAlaAsnArgGlnGlySerAlaArgAsnIleThrV 585
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```


1677 GCACAGCCTCAACTTCAGCAGCCGCGGCTCCGTGCGCAACTTGCTG 1726
585 aLysValGlnPheMetTyrGlyIAspProSerAsnAlaMetProVal 601
1727 TGGAGTGGACATGACAGGCGAGGAGCCGACGCGCTCTGCGGCTC 1776
602 IlePheGlyLysSerSerCysSerGlnPheSerLysGlnAlaTyrThrAl 618
1777 ACTTTGGCAAGTCCAGCTGACGATGATTACCCGCGAGCCTTCACACC 1826
618 aValValTyrHisAsnArgSerProAspPheHisGlnGlnIleLysVal 635
1827 GGGGGCTTACCAATACAAAGTCCCGAGTTCTACGAGGATTCAAGCTGC 1876
635 LysLeuProAlaThrLeuThrAspHisHisIleLeuPheThrPheTyr 651
1877 ACTTCCAGCCTGCGTGCAGAGAACATCACCCTGCTGTCCACTTTCAC 1926
652 HisValSerCysGlnGlnLysGlnAsnThrProLeuGlnThrProValG 668
1927 CATGTACGCTGCAGCCGCGCGGACACTGCGCTGAGACACCCGCTGG 1976
668 TyrThrThrPheLeuPheMetLeuGlnAsnGlyArgLeuLysThrGlyInP 685
1977 CTTTACTTGGATCCCACTGCTGACGACGGCGCTGAGAGACCGGCCCT 2026
685 heCysLeuProValSerLeuGlnLysProProGlnAlaTyrSerValLeu 701
2027 TCTGTCTCCAGTGTCTGTGGACACGCCGCCACGCTATTCCTGCTC 2076
702 SerProGlnValProLeuProGlyMetLysTyrValAspAsnHisLysG 718
2077 ACACCGGATGGCGCTCCGGCGCATGCGCTGGGTGAGCGTCAACAAGG 2126
718 ValAlaPheAsnValGlnValAlaValSerSerIleHisThrGlnAsp 735
2127 CGTGTCACTGTGGAGCTCACAGCCGCTGCTGTGTGACCCCGAGGACC 2176
735 rOtYrLeuAspLysPhePheAlaLeuValAsnAlaLeuAspGlnHisLeu 751
2177 CCACTTGGACAAATTCCTCACCCTGTGTCACGTCTGAGAGAGGAGGCC 2226
752 PheProValArgIleGlyAspMetArgIleMetGlnAsnAsnLeuGlnAs 768
2227 TTCCTATCCGCGCTCAAGACACTGTGTGAGCAGAGGCAACGTGGAGCA 2276
768 nGluLeuLysSerSerIleSerAlaLeuAsnSerSerGlnLeuGlnProV 785
2277 GAGCGTGGGGCGAGCTTGCAGACACTGCGCTGGCCAGCCCGCAACCC 2326
785 aLValArgPheLeuHisLeuLeuLeuAspLysLeuIleLeuLeuValIle 801
2327 TTTGGCTCTTCTCCACACAGTCTGTGACAGCTGCTGCTGTGCTCATC 2376
802 ArgProProValIleAlaGlyGlnIleValAsnLeuGlyGlnAlaSerPh 818
2377 AGGCCCCGATCATCAGTGGCCGATTGTGAACCTGGCCGTGGAGACCTT 2426
818 eGluAlaMetAlaSerIleIleAsnArgLeuHisLysAsnLeuGlnGlyA 835
2427 TGAAGCAATGGCCCATGTAGTACGCTTGTCAACGCGAGCCTGGAGGCG 2476
835 snHisAspGlnHisGlyArgAsnSerLeuLeuAlaSerTyrIleHisTyr 851
2477 CCGAGGATGCGCGGCTCACCTGCCACAGCTGCTGCTTACGTCCACTAC 2526
852 ValPheArgLeuProAsnThrTyrProAsnSerSerSerProGlyProG 868
2527 GCCTTTCGCTCTCTGCACTGAGCCAGCTCCGCGATGGGGCCCTC... 2574
868 yGlyLeuGlyLysSerValHisTyrAlaThrMetAlaArgSerAlaValA 885
2575CCAGTGCAGATGCAAGCTGCACACTGCGCGCTGGCTGTGTC 2617

885 rGProAlaSerLeuAsnLeuAsnArgSerArgSerLeuSerAsn 901
2618 GCCCGGAGGCTCTACCTGGCGGTTCCAAAGACATCACAGACAGCAAC 2667
902 ProAspIleSerGlyThrProThrSerProAspAspGlnValArgSerI 918
2668 CCTACCTCGCGCGGCGCTGCTGTGATGACAGGCTTCCCGCAT 2717
918 eIleGlySerLysGlyLeuAspArgSerAsnSerTyrValAsnThrGlyG 935
2718 CTTGGCCAGC..... 2727
935 LysProLysAlaAlaProTyrPglySerAsnProSerProSerAlaGluSer 951
2727 2727
952 ThrGlnAlaMetAspArgSerCysAsnArgMetSerSerHisThrGluTh 968
2727 2727
968 rSerSerPheLeuGlnThrLeuThrGlyArgLeuProThrLysLysLeuP 985
2728AAGCTGC 2734
985 heHisGlnGlnLeuAlaLeuGlnThrProValValCysSerGlySerValArg 1001
2735 TTCACAGAGAGCTGCTGTGACAGTGGGTGTCACAGACAGTCCGCTACG 2784
1002 GluSerAlaLeuGlnGlnAlaTyrPhePhePheGlnLeuMetValLysSe 1018
2785 GAGGCCATCTCCACAGCGCTGCTTCTTCCAGCTCATGTGGAGAAG 2834
1018 rMetValHisHisLeuTyrPheAsnAspLysLeuGlnAlaProArgLys 1035
2835 TATGGCTGACCTGCTGCTGTGGCCAGCAGTACAGACACCCCGCAAGC 2884
1035 eArgPheProGlnArgPheMetAspAspIleAlaAlaLeuAlaSerThr 1051
2885 TGGCTTCCCGGAGCGCTCTCTGACGACGACATCAGCTGTGTGGCTCT 2934
1052 IleAlaSerAspIleValSerArgPheGlnLysAspThrGlnMetValG 1068
2935 GTGGCGCTCGAGGTCATCACCCGCTGTCCACAAAGATGTGAGCGGCCGA 2984
1068 uArgLeuAsnThrSerLeuAlaPhePheLeuAsnAspLeuLeuSerValM 1085
2985 GCACCTCAACGCCAGCGCTGGCTTCTTCTCAGTGAACCTTCTGCTCTGG 3034
1085 eAspArgGlyPheValPheSerLeuIleLysSerCysTyrLysGlnVal 1101
3035 TGGACGGGGGCTTGTGTCTACGCTGTGCGCGGCCACTACAAACAGAGTG 3084
1102 SerSerLysLeuTyrSerLeuProAsnProSerValLeuValSerLeuAr 1118
3085 GCCACGCGGCTCCAGTCTCCCTAATCCAGACACCTGCTGAGACCTGCG 3134
1118 gLeuAspPheLeuArgIleIleCysSerHisGlnHisTyrValThrLeuA 1135
3135 CATGGAATTCACCGCATCTGTGACAGCACAGACACTGACTGACCTCA 3184
1135 snLeuProCysSerLeuLeuThrProProAlaSerProSerProSerVal 1151
3185 ACCTGCCCTGCTGCCCTGTACCTCTCAGCTGCGCTGCCCTCTCTGTG 3234
1152 SerSerAlaThrSerGlnSerSerGlyPheSerThrAsnValGlnAspG 1168
3235 TCTCTCCACACCTCCAGAGCTCCACCTTCCAGCCAAAGCCCGGAGCC 3284
1168 nLysIleAlaAsnMetPheGlnLeuSerValProPheArgGlnGlnHisT 1185
3285 CAAGGTGACGACGATGTTGAACTGAGTGAACATTCGCGGAGGAGCACT 3334

```

1185 ytleuAlaGlyLeuValLeuThrGluLeuAlaValIleLeuAspProasp 1201
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3335 TCCTAGCTGGCTCTGCTACGAGCTGCACTGGCCCTCGAAGCTGAG 3384
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1202 AlGluGlyLeuPheGlyLeuHisLysLysValIleAsnMetValHisas 1218
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3385 GCTGAAGGGGCTCTGCTGACACAAGAAGGCATCAGTGTGTGCACAG 3434
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1218 nleuLeuSerSerHisAspSerAspProArgTyrSerAspProGlnIleL 1235
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3435 CCTGCTATGTGGCATGACATGACATCCCGCTACGCCGAGGCCACTGTA 3484
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1235 yAlaArgValAlaMetLeuTyrLeuProLeuIleGlyIleIleMetGlu 1251
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3485 AGGCTGTGTGGCGGAGCTGTACCTGCCACTGCTTTCGATGTGACGGGAT 3534
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1252 ThrValProGlnLeuTyrAspPheThrGluThrHisAsnGlnArgLys 1268
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3535 ACCTTGCACAGGCTGCATGCTTGTGAGGGGCCAGTCAAGGCTCAG 3584
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1268 gProIleCysIleAlaThrAspAspTyrGluSerGluSer.....Glys 1283
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3585 ACTGGCCTCATGTCTGACTGACACACAGAGCGAAGGGGACATTGCGG 3634
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1283 erMetIleSerGlnThrValAlaMetAlaIleAlaGlyThrSerValPro 1299
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3635 GTACCATCAACCCCTGTGTGGCATGTGCTGTG6CCCCCTA... 3681
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1300 GlnLeuThrArgProGlySerPheLeuLeuThrSer.....Th 1312
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3682 .....GCCCTGTGCTCCGCGGACCATCTCCAGGGGCCACCAAC 3722
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1312 rSerGlyArgGlnHisThrPheSerIleGluSerAspSerLeu 1329
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3723 GGGCTTCCGCGAGGCTGTGCTGTCTGTGTGATGACCGGACCTTGC 3772
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1329 euIleCysLeuLeuThrValLeuLysAsnAlaAspGluThrValLeuGln 1345
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3773 TGGCGTGTGTGCTGTGGTGTGAAACACGAGCGGCGCTCGCGAG 3822
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1346 LysThrPheThrAspLeuSerValLeuGlnLeuAsnArgLeuLeuAspLe 1362
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3823 CGCTGGGCCACTGACCTGACACTCCCGCAGCTGGGACCTGTGTGACTT 3872
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1362 uLeuTyrLeuCysValSerCysPheGluTyrLysGlyLysValPheG 1379
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3873 GCTGTACTTTGGCTGGCTGCTTTGTAGTACAAAGGAAAGGCGCTTGG 3922
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1379 LuArgMetAsnSerLeuThrPheLysLysSerLysAspMetArgAlaLys 1395
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3923 AACGCATCAACGCCCTCACATTCAAAATCTCTGATATGAAGCGCGG 3972
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1396 LeuGluGluAlaIleLeuGlySerIleGlyAlaArgGlnMetValAr 1412
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3973 CTAGAGGAAGCCATTCTGGTACATCGAGCTCGACAAGAAATGATTGCG 4022
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1412 gArgSerArgGlyGlnLeuGlnArgSerProSerGlySerAlaPheGlys 1429
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4023 GCGAAGTCT.....GAGGAGGCGG.....TTTGGGA 4051
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1429 erGlnGlnAsnLeuArgTyrArgLysAspMetThrHisTyrArgGlnAsn 1445
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4052 ATCCAGAGAACGTGGCGTGGGGAAGAGGTGCACACACTGGAAGCAACC 4101
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1446 ThrGluLysLeuAspLysSerArgAlaGluIleGluHisGluAlaLeu 1462
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4102 TCAGACCGCGTGGACAAAGACCAAGTGAATGAAACACAGAGCCTTGGT 4151
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1462 eaSpGlyAsnLeuAlaThrGluAlaAsnLeuIleLeuAspThrLeuG 1479
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4152 GGAAGGAACCTGGCAACCGAGGCAAGCTAGTGTGTCGGACACATGG 4201
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1479 LuIleValValGlnThrValSerValThrGluSerLysGluSerIleLeu 1495

```

```

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4202 AGATATGTGTGACAGCGGTGATGCTTTCAGAAACCCGGAGAGGCTTTG 4251
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1496 GlyGlyValLeuLysValLeuLeuHisSerMetAlaCysAsnGlnSerAl 1512
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4252 GGGGAGGTGTGAAGGTTGTGCTGTACAGCCTGGGAGAGTGGCCAGAGTGC 4301
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1512 aValTyrLeuGlnHisCysPheAlaThrGlnArgAlaLeuValSerLysP 1529
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4302 CTTCTTGTGGACATGGCTGGGCACACCAGAGGGCCCTTGTGTCCAGT 4351
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1529 heProGluLeuLeuPheGluGluGluThrGluGlnCysAlaAspLeuCys 1545
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4352 TCCCGAGACTGCTGTTCAGAGAGACACAGGAGCTGTGCGGACCTGTGC 4401
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1546 LeuArgLeuLeuArgHisCysSerSerSerIleGlyThrIleArgSerH 1562
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4402 CTGAGGCTCTTACGACACTGTGGCAGCGCATTCAGACCATCCGACAGCA 4451
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1562 sProSerAlaSerLeuTyrLeuLeuMetArgGlnAsnPheGluIleGly 1579
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4452 CGCCAGCGCTGCTGTACCTGCTCATGCGACAGAACTTCAGATTCGGCC 4501
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1579 snAsnPheAlaArgValLysMetGlnValPrometSerLeuSerSerLeu 1595
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4502 ACACTTTGCCCTGTGTGAGATGCAAGTCAACATGTCTCTCTGCTGCTG 4551
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1596 ValGlyThrSerGlnAsnPheAsnGluGluPheLeuArgArgSerLeu 1612
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4552 GTGGGAGACGACGAACTTGTAGTGAAGAGCCTGGGACGTTTACTCA 4601
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1612 sThrIleLeuThrTyrAlaGluGluAspLeuGluLeuArgGluThrThrP 1629
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4602 AACCATCTCACTATGCTGTGAGAGACATGGGCTCGGAGACGACCT 4651
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1629 heProAspGlnValGlnAspLeuValPheAsnLeuHisMetIleLeuSer 1645
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4652 TCGCAGACACAGTCCAGAGACTGATGTTCAACCTGCACATGATCTTACG 4701
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1646 AspThrValLysMetLysGlnHisGlnGluAspProGluMetLeuLeas 1662
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4702 GACACGGTGAAGATGAAGAACACACAGAGAGACCTTGATGTCATGCA 4751
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1662 pLeuMetTyrArgIleAlaLysGlyTyrGlnThrSerProGlu...Arg 1678
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4752 CTTATGTACAGAAATGCGCGGGCTACAGGCTCACCGGACCTTGGCG 4801
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1678 euThrTyrLeuGlnAsnMetAlaGlyLysHisSerGluArgSerAsnHis 1694
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4802 TGACCTGTGTCACAAACATGGCCGGAACACAGCGGAGCTGGCAACAC 4851
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1695 AlaGluAlaAlaGlnCysLeuValHisSerAlaAlaLeuValAlaGlu 1711
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4852 GCCAGAGCGCCGACATGATGTGCACGCGCGCCCTGCTGTGGTGA 4901
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1711 rLeuSerMetLeuGlnAspArgLysTyrLeuProValGlyCysValThrP 1728
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4902 CTTGCGCTGCTCGAGGACCAAGCGCACCTGCGGTGGGTGCTTCT 4951
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1728 heGlnAsnIleSerSerAsnValLeuGluGluSerAlaValSerAsp 1744
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4952 TCCAGAACATCTCATCCACACTGCTAGAGAGTCCGCATCTCCGACAGC 5001
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1745 ValValSerProAspGluGluGlyIleCysSerGlyLysTyrPheThrG 1761
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5002 ATCTGTGCGCCGACGAGAGGGCTTCTGTCGGGAAGCACTTCACTGA 5051
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1761 uSerGlyLeuValGlyLeuGluGlnAlaAlaAlaSerPheSerMet 1778
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5052 GCTGGGCTGTAGGTTGCTGTGGAACAGGACCGGCTTACTTCAACATGG 5101
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1778 laGlyMetTyrGluAlaValAsnGluValTyrLysValLeuIleProIle 1794
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```


1396 uGIuGluAlaIleuGIySerIleGIyAlaArgGIuGIuMetValArga 1413
|||||
1702 TGAAGAACTATTCTTGGACATAGTGCCAGCAAGAAATGGTAGCGC 1751
1413 rGSerArgGIyGIuIleuGIuArgSerProSerGIySerAlaPheGIySer 1429
|||||
1752 GAAGCGGAGACGCTCGAAGAAACCATCTGGAAGTCCCTTGGGAAGT 1801
1430 GIuGIuAsnIleuArgTrpArgIysAspMetThrHisTrpArgGIuAsnTh 1446
|||||
1802 CAAGAAATTTTCAGCTGGAGAGAAAGATATGACTCACTGGCCTCAAAACAC 1851
1446 rGIuIysLeuAspIysSerArgAlaGIuIleGIuHisGIuAlaIleuIlea 1463
|||||
1852 AGAGAAGCTTGACAATACAGAGCAGAGATTGAACACGAAACACACTGATTG 1901
1463 sPGLyAsnIleuAlaThrGIuAlaAsnIleuIleIleuAspThrIleuGIu 1479
|||||
1902 ATGGAACCTGGCTACAGAAAGCAACCTAATCATTTTAGATACATTAAGAG 1951
1480 ILeValIleGIuThrValSerValThrGIuSerIysGIuSerIleuGIu 1496
|||||
1952 ATTGTGTTCACAGCCGTTCTCTTAACGGAATCCAAAGAGACATTCTTGG 2001
1496 yGIyValIleuIysValIleuIleuHisSerMetAlaCysAsnGIuSerAlaV 1513
|||||
2002 TGGAGTGCTAAAGAGTCTACTACACAGCATGGCCCTGAACCAAGTGCAG 2051
1513 aITyIleuGIuHisCysPheAlaThrGIuArgAlaIleuValSerIysPhe 1529
|||||
2052 TTTACTACAAACACTGTTTGTCTACACAGAGCCCTGTGTTCAAAAGTTT 2101
1530 ProGIuIleuIleuPheGIuGIuGIuThrGIuGIuIncIysAlaAspIleuCysIle 1546
|||||
2102 CCTGAACCTTATTTTGAAAGAAAGACAGACAGTCTGCTGATTATGGCT 2151
1546 uArgIleuIleuArgHisCysSerSerSerIleGIyThrIleArgSerHisP 1563
|||||
2152 CAGGCTTCTCCGACACTGTAGAGATGCGTACCAATACGCTCACACG 2201
1563 rOSerAlaSerIleuTyIleuIleuMetArgGIuAsnPheGIuIleGIyAsn 1579
|||||
2202 CCAGTGCCTCCCTTAACCTAATAGAGCAAAACCTTGAGATTGGGAAT 2251
1580 AsnPheAlaArgValIysMetGIuValPrometSerIleuSerSerIleuVal 1596
|||||
2252 AACTTTGCCAGGGTTAAATGCAAGTACATGATGCTATCTCTCTTGGT 2301
1596 IGIyThrSerGIuAsnPheAsnGIuGIuPheIleuArgArgSerIleuIysT 1613
|||||
2302 GGGACACATCTCAGAAATTTTAAATGAAGAATTTTAAAGACGTTCTTAAGA 2351
1613 hTrIleuThrTyIArgIuGIuAspIleuGIuIleuArgGIuThrThrPhe 1629
|||||
2352 CTATATTGACATATGCTGAAGAAGATCTGGAATTGAGGAAACAACATT 2401
1630 ProAspGIuValGIuAspIleuValPheAsnIleuHisMetIleuSerHis 1646
|||||
2402 CCTGATCCAGGTCCAGGATCTGTTTCAATCTCCATATGATCTTCTTGA 2451
1646 pThrValIysMetIysGIuHisGIuAspProGIuMetIleuIleAspL 1663
|||||
2452 TACTGTGAATAATGAAGAACACACAGAGGATCTGAAATGTTGATTGATC 2501
1663 eumetTyIArgIleuAlaIysGIyTyIArgIuThrSerProGIu...ArgIleu 1678
|||||
2502 TAATGTACGAATATGCCAAGGTTACCAAGACCTCTCCAGATCTCGGATGG 2551
1679 ThrTrpLeuGIuAsnMetAlaGIyIysHisSerGIuArgSerAsnHisAl 1695
|||||
2552 ACCTGTTGCAGACATGGCAGCAAGCACTACAGAAAGAAACATCATCC 2601
1695 aGIuAlaIleGIuIncIysIleuValHisSerAlaIleuValAlaGIuTyI 1712

|||||
2602 TGAAGCGCACAGTGTCTACTCCACTACAGACACTTGTTCGAATAT 2651
1712 eusermetIleuGIuAspArgIysTyIleuProValIGIyCysValThrPhe 1728
|||||
2652 TGAACATGCTGGAAGACCGGAATATCTTCTCTGTGGATGTGAACATT 2701
1729 GIuAsnIleSerSerAsnValIleuGIuGIuSerAlaValSerAspAspVa 1745
|||||
2702 CAGAAATTTTCATCTAATATGTTTAAAGAAATCTCGCGTCCAGATGATCT 2751
1745 IValSerProAspGIuGIuGIyIleCysSerGIyIysTyIArgPheThrGIu 1762
|||||
2752 GGTATCTCCAGATGAAGAAAGTATCTGCTCTGGAATACTTTACTAGAGT 2801
1762 eArgIleuValIGIyIleuIleuGIuAlaIleAlaIleSerPheSerMetAla 1778
|||||
2802 CAGGACTTGTGGATTAAGTGAACAAAGACAGCTCTTCTCTTANAGGCT 2851
1779 GIyMetTyIArgIuAlaValAsnGIuValTyIysValIleuIleProIleH 1795
|||||
2852 GGCAATGATGAGACAGTATTAAGATTACAAAGTACTTATCTCATTTCA 2901
1795 sGIuAlaAsnArgAspAlaIysIysIleuSerThrIleHisGIyIysIleu 1812
|||||
2902 TGAAGCTAATCGGATGCCAAAGAACTATCCACAATTCATGATGAACCTTC 2951
1812 IngIuAlaPheSerIysIleValHisGIuSerThrGIyTrpGIuArgMet 1828
|||||
2952 AAGAAAGATTTCAGCAAAATTTGTCATCAGAGTACTGGCTGGAGCGGATG 3001
1829 PheGIyThrTyIArgValIGIyPheTyIArgIyPheIysPheIysPhe 1845
|||||
3002 TTTGGACACTATTTTGTGTTGTTTATGAAACAAAGTTCGGGATTT 3051
1845 uAspGIuGIuGIuPheValTyIysGIuProAlaIleThrIysIleuAlaG 1862
|||||
3052 GGATGAACAAGAAATTTGTTTACAAAGAGCCCTGCATTAACCAACTTGCAG 3101
1862 IuIleSerHisArgIleuGIuGIyPheTyIArgIuArgPheGIyIysIleuAsp 1878
|||||
3102 AGATATCTCACAGATGTGAGGATTTTACGAGCAAGATTTGAGAGAGAT 3151
1879 ValValGIuValIleIysAspSerAsnProValAspIysCysIysIleuAs 1895
|||||
3152 GTGGTTGAAGTAATCAAAAGCTTAATCTGTGACAAAGTGAATAGA 3201
1895 pProAsnIysAlaTyIleGIuIleThrTyIArgValGIuProTyIArgPheAspT 1912
|||||
3202 TCTTAACAGGCAATATTTAGATTACCTATGTGCAACCATCTTGTGACA 3251
1912 hTrIyGIuMetIysAspArgIleThrTyIArgPheAspIysAsnTyIAsnIleu 1928
|||||
3252 CATATGAGATGAGAGACAGAAATTCACCTATTTCGACAAAAATTCATCTT 3301
1929 ArgArgPheMetTyIysCysThrProPheThrIleuAspGIyArgAlaHisGI 1945
|||||
3302 CGTGATTCATGATGTAAGTACACCTTACTTATGATGGCGTGCCATGG 3351
1945 yGIuIleuHisGIuIlePheIysArgIysThrIleuThrThrSerHisA 1962
|||||
3352 GGAACCTCATGAAACAAATTCAAAAGAGACCATCTGACACGCTCATG 3401
1962 IArgPheProTyIleIysThrArgValAsnValThrHisIysGIuGIuIle 1978
|||||
3402 CCTTCTCTATATTAACAAGGCTCAATGCTCAATTAAGAAGAGATC 3451
1979 IleuThrProIleGIuValAlaIleGIuAspMetGIuIysIysThrGI 1995
|||||
3452 ATCTTAACACCAATTTGAAGTGTATATGAGACATGCAAGAAAGACACA 3501
1995 nGIuIleuAlaPheAlaThrHisGIuAspProAlaAspProIysMetIleuG 2012
|||||

```

3502 GGAGTTGGCATTTGCACACATCAGATCCCGGACAGCCCAAAATGCTTC 3551
2012 lmeValLeuGInGlySerValGlyThrThrValAsnGInGlyProLeu 2028
3552 AGATGGTACTCCAGGAGTCTGTAGGACACACATGATCAGGGGCTTTG 3601
2029 GluValAlaGlnValPheLeuSerGluIleProSerAspProLysLeuH 2045
3602 GAAGTGGCCAGGTTTTTCTGTGAAATACCTAGGACCCAAAGGCTCTT 3651
2045 eARGHSHLSAsnLysLeuArgLeuGlySerPheLysAspPheThrLysArgC 2062
3652 CACACATCAATTAATAACTGGAGACTCGCTTTAAAGATTTTACTAAAGGT 3701
2062 ySLuAspAlaLeuArgLysAsnLysSerLeuIleGlyProValGlnLys 2078
3702 GTGAAATGCTCTTAAGAAAAATAAAGCTTAATTGGCCGGCATCAAAAG 3751
2079 GluTyrGlnArgGluLeu..GlyLysLeuSerSerPro 2090
3752 GAGTATCAAAAGGAGACTGGAGAGAACTATCATCGCT 3789

seq_name: gb_pat:AX172980

seq_documentation_block:
LOCUS AX172980 4200 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 107 from Patent WO0142297.
ACCESSION AX172980
VERSION AX172980.1 GI:14597958
KEYWORDS
SOURCE
ORGANISM
.
synthetic construct.
synthetic construct
artificial sequence.
1 (bases 1 to 4200).
REFERENCE
1 Lu,P., Garman,J.D. and Candia,A.F.
AUTHORS
Clasp-3 transmembrane protein
TITLE
Patent: WO 0142297-A 107 14-JUN-2001;
JOURNAL
Arbor Vita Corporation (us)
FEATURES
location/Qualifiers
source 1..4200
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Polynucleotide fragment"
BASE COUNT 1262 a 844 c 907 g 1187 t
ORIGIN
alignment_scores:
Quality: 6171.50 Length: 1226
Ratio: 5.088 Gaps: 2
Percent Similarity: 98.940 Percent Identity: 98.206

alignment_block:
US-09-737-246-2 x AX172980 ..
Align seg 1/1 to: AX172980 from: 1 to: 4200

866 GYPRPGYGLYLEUGLYLSERVALIHSTRALATHMETALARGSE 882
|||||
3 GGGCCCGGCTCCGCTCGGTGACGGATGCAATGATTCGAAATTC 52
882 rAlaValArgProAlaSerLeuAsnLeuAsnArgSerArgSerLeuSerA 899
|||||
53 GGCA.....CGAGTAAATGCTTCTGCAACCTTAGTA 84
899 snSerAsnProAspIleSerGlyThrProThrSerProAspGluVal 915
|||||
85 ATAGCAATCCAGATATATCTGGAGCTCCACAGCTACACAGATGATGAAGTT 134
916 ArgSerIleIleGlySerLysGlyLeuAspArgSerAsnSerTrpValAs 932
|||||
135 GATCAATCATCTGGAGTAAAGGTTAGTTCCTCCAAATCTCTGGGTTAA 184

```

932	ntnrglygylprolyalaalaprotpgilyseransproserprosera	949
185	cactggggggtccaaagactggcccatggggatccaaacccactgcaatg	234
949	laqluserrhnginalametaspargsercysanargmetserberhs	965
235	cagaaatcaacacagggatgagatcgaaagtggaaatcgatgcttgcgac	284
966	thngluthrserserpheuglnthrleuthrnlargyleprothly	982
285	acagagacgctcaggttcttacaacaattaaacggagccttaacacactaa	334
982	slysleuphehisglgluleualauleuglntrpyalvalcysserglys	999
335	aaagctttttcacagagacctggctttcgatgggttggtttgacatgsc	384
999	etvalargjuseralaleuenglinalatprpherphegileuemet	1015
385	ggcgtggggatagctttggcaaacgctgggttttttgaaattaatg	434
1016	vallyssermetvalhishisleuylrphenaasplysleuglala	1032
435	gtmaagacatgggtgacacattatnacttaattgataaacttgagacgtc	484
1032	oarglysserargpheprogluargpheketaspsllalealaleu	1049
485	aaagaaaaagctgtttccgaacgctttacgatgacattgcagctcttg	534
1049	alserthrileaserrapllevalserargphglnllyasphrctlu	1065
535	tccagcagatgctatgatatnagtttcacgatttcacaaagacacagaa	584
1066	metvalgluargleuasnhrserleualaphelaleuasasplleu	1082
585	atggttgagagatcatatnagaacgctgcattctttcattgatgactgt	634
1082	uservalmetasparglyphevalpheserleulleysercysyl	1099
635	gctgttgatgagacagacgattgttttgacctgtttaaagctcgctata	684
1099	ysglinalserserlyseuylrserleuproasnproservalleu	1115
685	aaagagctgttcaagcgtttactatnaccgaatccagagctgtcggtg	734
1116	serleuargheuasphelenuarqlllecysserhishislytyva	1132
735	tccctgagcctgattttttagaatcatctgacactatgacactatagt	784
1132	lthrleuasnleuprocysserleuethrprofolaserproserp	1149
785	tacatttaacttaacctgcagcgttacttaactccacctgcattccatc	834
1149	roseralserseralatrhrserglnserserglypheserthasnyal	1165
835	cttctgtttcttctgcacacatgtcagagcttgcgatttttcagacaa	884
1166	glnaspglnlysillealasnmetphegileuenservalprophear	1182
885	caagaccamaagatttccaaatnctgttgataattatccgctttccgcga	934
1182	nglnhistrylleualaglyleuvalleuthrcluleualavalilleua	1199
935	acagcattttttggcagcagcctgtgttaacagagctgctgctat	984
1199	sprprokspralagluclyleuphegilyleuhslyslsvalilleasmet	1215
985	accctgatctctaaagacactgtttgattgcataaagaaagtatcaaatg	1034
1216	valhishaspleuenserserhishspserspproarglysserap	1232
1035	gtacacaaattttactctccagctagcactcaaacccggcgatctgtgacc	1084
1232	oglnlileylsalaargvalalameleuetyrleuproleuileglyl	1249

1085	TCAGATGAAAGGCTGAGAGGCCAATGGTGTATCTACCTCATATGGTATTA	1134
1249	lmeGcltHrVtAlProGlnLeuYrAspPheHrGluTrHnIsAsnGln	1265
1135	TCATGAAACGTCTACCTGAGCTGTATGATTTTACAGAACTCACATCAA	1184
1266	ArgGlyArGProIleCysIleAlaTrAspApTyGlsuSerGlnSerG	1282
1185	CGAGGAGACCAATTTGTATAGCCATGATGATTAATAAGTAGAGAGCG	1234
1282	yserHctIleserGlnTrValAlaMetAlaIleAlaGlyTrHnSerValP	1299
1235	AAGTATGATTAACCCAGACCGTTGCCATGCGCAATGCGAGGACATCGGTC	1284
1299	roGlnLeuHrArGProGlySerPheLeuLeuTrHnSerTrHnSerGlyArg	1315
1285	CTCACTAAACAAGGCGCTGGCAGTTTCTCTCCACGTCACAGATGGCAGG	1334
1316	GlnHnIsHrTrHnPheSerAlaGluSerSerArGySerLeuLeuIleCysIle	1332
1335	CAACACACTACTCTTTTCAGAGAAATCAAGTCGAGGCGCTTTTGATCTGTCT	1384
1332	uLeuTrpValLeuLysAsnAlaAspGluTrHnValLeuGlnLysTrpHet	1349
1385	ACTTTGGGTTCCAAAAATGCAAGATGAACAGTTCTCAAGAAAGTGTTTA	1434
1349	hrAspLeuSerValLeuGlnLeuAsnArGleuLeuAspLeuLeuTyTrLeu	1365
1435	CAGATCTCTCACTCTTGCAAGCTTAACCGGCTATTAAGATCTGCTTTATCTC	1484
1366	CysValSerCysPheGluTyTrLysGluTyLysValPheGluArGmetAs	1382
1485	TGHTGTCTTGGCTTTCAGATTAAGGGAAAAAGTGTTCAGCAATGA	1534
1382	nsertLeuTrHnPhelyLysSerLysAspMetArGAlaLysLeuGlnGlu	1399
1535	TAGCTTGACCTTTAAGAAATCAAAAAGACATGACGCAACGCTTGAAACAG	1584
1399	IaIleLeuGlySerIleGlyAlaArGlnGlnMetValArGArgSerArg	1415
1585	CTATTCCTTGGAGCGCTAGGTGCCAGGCAAGAAATGTCATCGGCAAGCCGA	1634
1416	GlyGlnLeuGluArGSerProSerGlySerAlaPheGlySerGlnGluAs	1432
1635	GGACACCTCGAGAGAAGCCATCTGGAAAGTCCTTTGGAAAGTCAAGAAA	1684
1432	nLeuArGTrpArGlyAspMetTrHnIsTrpArgLhAsnTrGluLysL	1449
1685	TTTGAGAGTGAAGAAAGATATGACTCACTGCGCTCAAAACACAGAGAGGC	1734
1449	eulaPyrSerTrpGluGluIleGlnHnIsGluAlaLeuIleAspGlyAsn	1465
1735	TTGACAAATCAAGACAGAGATGTGAACCGAAGGACGATTTGATGAGAAAC	1784
1466	LeuAlaTrHnGluAlaAsnLeuIleIleLeuAspHrLeuGlnIleVala	1482
1785	CTGGCTACAGAAAGCAACCTAATCATTTTAAATATCATTAAGATGTGTGT	1834
1482	IglInTrValSerValTrHnGluSerLysGlnSerIleLeuGlyGlyValL	1499
1835	TCAGACCGGTTTCTGTAAACGGAATCCAAAGAGACGATCTTGTGTGAGATGC	1884
1499	eulysValLeuLeuHnIsSerMetAlaCysAsnGlnSerAlaValTyTrLeu	1515
1885	TAAAAATGGTACTACACACACAGGCGCTGTAAACAAAGTGCAGTTTATCTA	1934
1516	GlnHnIsCysPheAlaTrHnGlnArGluAlaLeuValSerLysPheProGlu	1532
1935	CAACACTGTTTGTCTACACAGAGAGCCTTGGTTTCAAAAGTTTCTGTAGACT	1984
1532	uLeuPheGluGlnGluTrHnGlnGlnCysAlaAspLeuCysLeuArGLeuL	1549

1985	CTTAATTGGAAGAAGACAGACAGACAGTGTGTCGATTATTAAGCTCCAGGCTTC	2034
1549	euatrgHsCysSerSerSerIleGlyThrIleArgSerHisProSerAla	1565
2035	TCCGACACGTGAGCAGTACCATTCGGACAAATACGGTGCACACCCAGTACC	2084
1566	SerLeuTyrlLeuLeuMetArgGlnAsnPheGluIleGlyAsnAsnPheAl	1582
2085	TCCCTTACCCTACTAATATGAGCCAAACCTTTGAGATGGGAATTAACCTTTGC	2134
1582	aarvValLysMetGlnValProMetSerLeuSerSerLeuValGlyHis	1599
2135	CAGGGTTAAATACAGAGTCCCATATGCTATATCCCTCTGTGGGACAT	2184
1599	ercGlnAsnPheAsnGlnGluPheLeuArgArgSerLeuLysThrIleLeu	1615
2185	CTCAGAAATTTTATATGAGAAATCTTAAGACGTTCTTAAGACATATATG	2234
1616	ThrTyrlAlaGlnGlnAsnLeuGlnLeuLeuArgGlnThrThrPheProAspG	1632
2235	ACATATGCTGAGAAAGATCTCGAAATTTGAGGAAACAAACATTTTCCGTATCA	2284
1632	nValGlnAsnLeuValPheAsnLeuHisMetIleLeuSerAspThrValL	1649
2285	GGCTCCAGGATCTGGTTTAAATCTCATATGATTTCTGTATCTGTGA	2334
1649	ysMetLysGlnHisGlnGlnAsnProGlnMetLeuIleAspLeuMetTy	1665
2335	AAATGGAAGACACCCAGAGAGATCTCGAAATGTATATGATCTAATGTAC	2384
1666	ArgIleAlaLysGlyTyrlGlnThrSerProGlu...ArgLeuThrTyrLe	1681
2385	AGAAATGCCAAAGGTTATCCACAGACCTCTCCAGATCTGCAGTTGACCTGGTT	2434
1681	uGlnAsnMetAlaGlyLysHisSerGlnArgSerHisAlaGlnAlaA	1698
2435	GCGAACAATGCGAGGACACTCCAGAACGAAGCATATGCTGAAGCTG	2484
1698	IacGlnCysLeuValHisSerAlaAlaLeuValAlaGlyTyrLeuSerMet	1714
2485	CACAGTGTCTAGTCCATCAGCAGCACTGTGTGCGAATTTTGAGCATG	2534
1715	LeuGlnAspArgLysTyrLeuProValGlyCysValThrPheGlnAsnI	1731
2535	CTGGAGAGACCGAAATATCTCTCTGCGGATGTGTAACATTTCAAGATAT	2584
1731	eSerSerAsnValLeuGlnGlnSerAlaValSerAspAspValValSerP	1748
2585	TTTCATCTAAATGTTTGAAGAAATCTGCGGCTCTCAATCATGTGGTATCTC	2634
1748	roAspGlnGlnGlyIleCysSerGlyLysTyrPheThrGlnSerCylLeu	1764
2635	CAGATGGAAGAGTATCTGCTCTGGAAATATCTTACTGAGTACAGGACTT	2684
1765	ValGlyLeuLeuGlnValAlaAlaAlaSerPheSerMetAlaGlyMetTy	1781
2685	GTCGGATTAATGGAACAACAGAGTGTCTCTCTATAGCGTCGAGTATGA	2734
1781	rgLualAlaValAsnGlnValTyrLysValLeuIleProIleHisGlnAlaA	1798
2735	TGAAGCAGTTAAAGAGATTACAAATACTTATTTCTTATTCATCAAGCTA	2784
1798	snArgAspAlaLysLysLeuSerThrIleHisGlyLysLeuGlnGlnAla	1814
2785	ATCGGGATGGAAGAAACATATCCACATTTCAAGTAACTTCACAAACCA	2834
1815	PheSerLysLleValHisGlnSerThrGlyTyrPGLuArgMetPheGlyT	1831
2835	TTTCAGCAAAATTTCTTCATCAGATACTGCGTCGGACCGGATCTTTGGCAC	2884
1831	TyrTyrlPheArgValGlyPheTyrGlyThrLysPheGlyAspLeuAsnGlnG	1848
2885	CTATTTTTCGTTGGTTTATATGAAACCAAGTTCGGGAGTTTGTGATGTAC	2934

```

1848  InGluPheValTyrLysGluProAlaIleThrLysLeuAlaIleSer 1864
1849  |||||
2935  AAGAAATTGTTTACAGAGAGCCGCAATACCAAACTTCCAGACATATCT 2984
1865  HisArgLeuGluGlyPheTyrGlyGluArgPheGlyGluAspValAlaG1 1881
2985  CACAGATTGGAGGATTTCAGGAGAAAGATTGGAGAGATGTGGTGA 3034
1881  UValIleLysAspSerAsnProValAspLysLysLeuAspProAsnL 1898
3035  AGTAATCAAGACTCTAATCTGTAGACAAAGTGAATTAAGATCTAACA 3084
1898  ySalTyrIleGlnIleThrTyrValGluProTyrPheAspThrTyrGlu 1914
3085  AGGCATATTTGCGATTACCTATGTGAGCCATCTTCTACACATATGAG 3134
1915  MetLysAspArgIleThrTyrPheAspLysAsnTyrAsnLeuArgPhe 1931
3135  ATGAAAGACAGAAATCACCTATTTCGACAAAAATTAACAATCTTCGAT 3184
1931  eMetLysThrProPheThrLeuAspGlyArgAlaHisGlyGluLeuH 1948
3185  CATGTACTGTACACCTTACTTGTAGATGGCGTCCCATGGGAACTTC 3234
1948  IsGluGlnPheLysArgLysThrIleLeuThrThrSerHisAlaPhePro 1964
3235  ATGAACAATTCACAAAGAGAACCATTTCTGACTACGCTCATCTTCTTC 3284
1965  TyrIleLysThrArgValAsnValThrHisLysGluGluIleIleLeuTh 1981
3285  TATATTAACAAGAGGCTCAATGTCATCTCAATAAGAAAGATCATCTTAA 3334
1981  rProIleGluValAlaIleGluAspMetGlnLysLysThrGlnGluLeuA 1998
3335  ACCAAATTGAAGTTGCTATTGAGGACATGCAGAAAGACACAGAGATTGG 3384
1998  lAPheAlaThrHisGlnAspProAlaAspProLysMetLeuGlnMetVal 2014
3385  CATTTGCACACATCAGATCCCGCAGACCCCAAAATGCTTCAGATGGTA 3434
2015  LeuGlnGlySerValGlyThrThrValAsnGlnGlyProLeuGluValAl 2031
3435  CTCACAGGATCTGTAGGACACACAGTGAATCAGGGGCTTTCGAGATTGC 3484
2031  aGlnValPheLeuSerGluIleProSerAspProLysLeuPheArgHisH 2048
3485  CCAAGTTTTCGTCTGAATACCTAGTACCCCAAGCTCTTCAGACATC 3534
2048  lAsnLysLeuArgLeuGlySerPheLysAspPheThrLysArgGlyAsp 2064
3535  ATTAATAACTGCGACTGCTGTTAAAGATTTTACTAAAGAGTGTCAACAT 3584
2065  AlAlaLeuArgLysAsnLysSerLeuIleGlyProValGlnLysGluTyrG1 2081
3585  GCCTTAAGAAAAAATAGAGCTTAATTTGGCGGTTCAAAAGAGATATCA 3634
2081  nArgGluLeuGlyLysLeuSerSerPro 2090
3635  AAGGGAATTGGGAAACTATCTTCGCT 3662
seq_name: gb_pat: AX172880
seq_documentation_block:
LOCUS AX172880 4144 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 7 from Patent W00142297.
ACCESSION AX172880
VERSION AX172880.1 GI:14597915
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

```

REFERENCE 1 (bases 1 to 4144)
AUTHORS Lu, P., Garman, J.D., and Candia, A.F.
TITLE CLASP-3 transmembrane protein
JOURNAL Patent: WO 0142297-A 7 14-JUN-2001;
Arbor Vita Corporation (US)
FEATURES
source
1. 4144
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="preliminary CLASP-3 cDNA"
<2..3610
/note="unnamed protein product"
/codon_start=1
/db_xref="GI:14597916"
/translation="RVNRSRLSNPNISGTPSPDDEVRSTIGSKGLDRNSWNT
GGKAPMGWNSPSESTQAMDRCNRMSSHTETSFLOTGLPKLHELEAL
QWVCSGVRESALQAMFFELMWKSVNHLLEAKRSEPERFMDIALVAL
STIASDIVRFQKDEVREBLTSLAFPLNDLSYMDGFVESLTKSCYKQVSRLYS
LNPVSIVSLRDLFRITCSHRYVTLPGLSLTPPASPSVSSATSSGGSFVY
DQKATANNEELSVPRQHYTAGLVLTSLAVILDPDAGLGLKAYINMHNLSH
DSDPRTSDQIKARVAMLYPLIGLIMETVQLDFTFHNQGRPTLATDYSES
GSMISQTVAMALAGTSVQLTRPGSFLLTSGRHTFSAESSKSLICLAWLKA
DETVALQKFTDLSTVLQRLDLILCYSCFEYKQKVFEMNSLTFFKSKDMAKLE
EALISIGARQEMVRSRGOLERSPGSAGSOELRWRKDMTHRMOWTEKLSRAE
TEHEALINGNLATEANLIIIDTLEIVQTVYTESKEIIGGLVKILLHSAQNSAY
YIQHCFATQALVSKFPELLEPEETEOCADICRLIRHCSSIGTIRSPASLYLM
RQNFETIGNFARVAKQVPMSSLSLVGTSQNFNEETLRSLKTLITLYAEDELLETFE
PDQVDLVFNLMILSDTVKMKHEQDEPMLIDMYRIANKQVSPDLRLMLQNMG
KHSERSNHAEOCLVHSAALVAEYLSMLDEPKYLPVCCVTFQNISVLEESAVD
VSPDEEGISGKFTYESGLVLEQAAASFMSMGMYPAVNEVYKVLIPHEARNDAK
KSTTHGKLOEAFSKIVHOSGMERMFPQTVRGVGYKRGDLDQEQEVKPEATIKL
AEISRLGFTYGERGEDVEYIKRNSNYDKCKDPKAYIQITVEYEPPTYEKDR
ITTFDKNTNLRREMTCTPPTLDGRHAGLHDFQRKTLITTSHPAYTKTENVATHE
ELIIPLEVAIEDMQKTPQELAFATHODPAPKMLQWLVQSVGTYNQGLEVAQVF
LSEIPSDPKLFENHNKRLCFKDFPKCEDALRRNKLISGVQKEYORELCKLSP"
BASE COUNT 1252 a 829 c 890 g 1173 t
ORIGIN
alignment_scores:
Quality: 6157.50 Length: 1201
Ratio: 5.131 Gaps: 1
Percent Similarity: 99.917 Percent Identity: 99.750
alignment_block:
US-09-737-246-2 x AX172880 ..
Align seq 1/1 to: AX172880 from: 1 to: 4144
891 LeuSnaArgSerArgSerLeuSerAsnSerAsnProAspIleSerGlyTh 907
:::|||||
5 GTAATCGTTCTCGAAGCCTTAGTAATAGCAATCCAGATATATCTGGAC 54
907 rProThrSerProAspAspGluValArgSerIleIleGlySerLysGlyL 924
55 TCCACGCTCACAGATATGAAGTTCATCAATCATCGGAGATGAAGGTT 104
924 eAspArgSerAsnSerTrpValAsnThrGlyGlyProLysAlaAlaPro 940
105 TAGATCGCTCCCAATTCCTGGGTTAACACTGTGTGTCACAAAGCTGC 154
941 TrpGlySerAsnProSerProSerAlaGluSerThrGlnAlaMetSpar 957
155 TGGGGATCCAAACCCAGTCCAGAGTCAACACAGGCTATGATGATG 204
957 gSerCyAsnArgMetSerSerHisThrGluThrSerSerPheLeuGlnT 974
205 AAGTTGTAATCGTATGTCTTCGACACACAGAGAGCTCAAGTTTCTTCA 254
974 hIleuThrGlyArgLeuProThrLysLysLeuPheHisGluLeuAla 990
255 CATTAACGGGACGCTTACCACTAAAAAGCTTTTCACAGAGAGCTGGCT 304

```


991 LeuGlnTrpValIValCysSerGlySerValArgGluSerAlaLeuGlnG1 1007
|||||
305 TTGCAGTGGGTTGTTGCAGTGGCAGCCTTCGGAAATACGCTTTGCCAAC354
1007 nAlaTrpPhePheGluLeuMetValIlysSerMetValHisHisLeuT 1024
|||||
355 AGCGTGGTCTTTTGAATTAATGTAAGAGCATGCTGCACCATTTAT 404
1024 yTrPheAsnAspLysLeuGluAlaProArgLysSerArgPheProGluArg 1040
|||||
405 ACTTTAAATGATAAAGCTGAGGCTCCAAAGAAAGTCGTTTCCAGAACGT 454
1041 PheMetAspAspIleAlaAlaLeuValSerThrIleAlaSerAspIleVal 1057
|||||
455 TTCAATGAGTGAATTCGACGCTTGTGCACGAGATGCTGATGATATAGT 504
1057 LSerArgPheGlnLysAspTrpGluMetValGluArgLeuAsnThrSerL 1074
|||||
505 TTCACGATTTTCAGAGGACACAGAAATGTTGAGAGACTCAATACCAAGCC 554
1074 euAlaTrpPheLeuAsnAspLeuLeuSerValMetAspArgGlyPheVal 1090
|||||
555 TTGCATCTCTTCTCAATATCTGTTGTGTATGACAGACAGAGATTTGTT 604
1091 PheSerLeuIleLysSerCysTrpLysGlnValSerSerLysLeuTrpSe 1107
|||||
605 TTTACGCTTATTAAGTCCTGCTATTAACAGGGTCTTCAAAAGCTTTACTC 654
1107 rLeuProAsnProSerValIleuValSerLeuArgLeuAspPheLeuArgI 1124
|||||
655 ATTACCGAATCCAGTGTCTGTGTCCTTGAGCGCTGATTTTTCACGAA 704
1124 leIleCysSerHisGlnHisTrpValIThrLeuAsnLeuProCysSerLeu 1140
|||||
705 TCATCTGAGTCATGACGACTATGTATTAACCTTAACCTCGCAGCTTA 754
1141 LeuThrProAlaSerProSerProSerValSerSerAlaThrSerG1 1157
|||||
755 CTTACTCCACCTGCATCTCCATCCATCCTCTGTCTTCTGCACACTTCA 804
1157 nSerSerGlyPheSerThrAsnValGlnAspGlnLysIleAlaAsnMetP 1174
|||||
805 GAGTCTGTGATTTTCTACGAATGTACAAAGACCAAAAGATTGCCAATATGT 854
1174 heGluLeuSerValIProPheArgGlnGlnHisTrpLeuAlaGlyLeuVal 1190
|||||
855 TTGAATTTATCCGTGCCCTTCCGCCAACAGCATTTATTTGGCAGGACTTGTG 904
1191 LeuThrGluLeuAlaValIleLeuAspProAspAlaGluGlyLeuPheG1 1207
|||||
905 TTACACAGAGCTGGCTGCTCATTTTAAACCTGATGCTGAAGGACTTTGG 954
1207 yLeuHisLysLysValIleAsnMetValHisAsnLeuLeuSerSerHisA 1224
|||||
955 ATTGCATAAGAAAGTCATCAATATATGTATACACAAATTACTCTCCAGTCA 1004
1224 sPheSerAspProArgTrpSerAspProGlnIleLysAlaArgValAlaMet 1240
|||||
1005 ACTCAGACCCCGCGGTACTCTGACCTCAGATTAAGGCTCGAGTGGCCATG 1054
1241 LeuTrpLeuProLeuIleGlyIleIleMetGluTrpValIProGlnLeuTy 1257
|||||
1055 TTGTATCTATCCTCTGATTTGATTTATCATGAAACTGTACCTCAGACTGTA 1104
1257 rAspPheTrpGluTrpHisAsnGlnArgGlyArgProIleCysIleAlaT 1274
|||||
1105 TGATTTTACAGAACTCACAATCAACAGAGAACCAATTTGTATAGCCA 1154
1274 hrAspAspTrpArgLysSerGlySerMetIleSerGlnThrValAla 1290
|||||
1155 CTGATGATTTATGAAGTGAGAGCGGAAGTATGATTAAGCCAGACCGTTGCC 1204

1291 MeAlaIleAlaGlyThrSerValProGlnLeuThrArgProGlySerPh 1307
|||||
1205 ATGGCAATGCCAGGACATGCGTCCCTCAACTAACAAGGCTTGCGAGTTT 1254
1307 eLeuLeuThrSerThrSerGlyArgGlnHisThrThrPheSerAlaGluS 1324
|||||
1255 CCTCCATCAGCTCAACAGAGTGGCAGGCAACACTACTACTTTTCAGCAGAA 1304
1324 eSerArgSerLeuLeuIleCysLeuLeuTrpValLeuLysAsnAlaAsp 1340
|||||
1305 CAAGTCGAAGCCTTTGATCTGTACTCTTGGGTCTCCAAAAATGCAAT 1354
1341 GluThrValLeuGlnLysTrpPheThrAspLeuSerValLeuGlnLeuAs 1357
|||||
1355 GAAACATTTCTACAGAAAGTGGTTACAGATCTCTCAGTCTTGACGCTTAA 1404
1357 nArgLeuLeuAspLeuLeuTrpLeuCysValSerCysPheGluTrpLysG 1374
|||||
1405 CCGGCTATTAGATCTGCTTATCTGTGTCTGTCTTCTTGAAGTATAAAG 1454
1374 LysLysValIlePheGluArgMetAsnSerLeuThrPheLysSerLys 1390
|||||
1455 GGAATAAAGTTTGAAAGATGAATAGCTTGACCTTTAAGAAATCAAAA 1504
1391 AspMetArgAlaLysLeuGlnGluAlaIleLeuGlySerIleGlyAlaAr 1407
|||||
1505 GACATGAGACCAAAAGCTTGAAGAAGCTATCTTGGAGAGCTTAGTCCAG 1554
1407 gGlnGluMetValArgArgSerArgGlyGlnLeuGluArgSerProSerG 1424
|||||
1555 GCAAGAAATGTACCGCGGAACCGAGAGACGTCGAGAGAACCCCATCTG 1604
1424 LysSerAlaPheGlySerGlnGluAsnLeuArgTrpArgLysAspMetThr 1440
|||||
1605 GAACTGCCCTTTGGAGTCAAGAAATTTGAGGTGAGAGAAAGTATGACT 1654
1441 HisTrpArgGlnAsnThrGluLysLeuAspLysSerArgAlaGluIleG1 1457
|||||
1655 CACTGGCGTCAAAAACACAGAAAGCTTGACAAATCAAGACAGCAAGTTGA 1704
1457 uHisGluAlaLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeuIleT 1474
|||||
1705 ACAGGAAGCACTGATGATGAAACCTGGCTACAGAAAGCAAACTATATCA 1754
1474 LeuLeuAspThrLeuGluIleValValGlnThrValSerValIThrGluSer 1490
|||||
1755 TTTTAGATACATTAGAGATTGTTGTGAGACCGTTCGTGAAGGATGCC 1804
1491 LysGluSerIleLeuGlyGlyValLeuLysValLeuLeuHisSerMetAl 1507
|||||
1805 AAAGAGAGCATTTCTGTGTGAGAGTGTAAAGTGTACTACACAGCATGGC 1854
1507 aCysAsnGlnSerAlaValIleTrpLeuGlnHisCysPheAlaThrGlnArg 1524
|||||
1855 CTGTAAACCAAAAGTGCATTTATCTACAAACACTGTTTGTCTACACAGAG 1904
1524 IAlaLeuValSerLysPheProGluLeuLeuPheGluGluGluTrpGluGln 1540
|||||
1905 CTTTGTTTCAAAAGTTTCTGAACTCTTATTTGAAGAGAGCAAGAGCAG 1954
1541 CysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSerIleG1 1557
|||||
1955 TGTCCTGATTTATGCCACAGGCTTCTCCGACACTGATGAGATGACATCGG 2004
1557 yThrIleArgSerHisProSerAlaSerLeuTrpLeuLeuMetArgGlnA 1574
|||||
2005 TACATATCGGTCACACCCAGTGCCTCTTCTTACTATATAGAGGAAA 2054
1574 snPheGluIleGlyAsnAsnPheAlaArgValLysMetGlnValPromet 1590
|||||
2055 ACTTTGAGATTGGGAATTAACCTTGGCCAGGGGTAAAGTCAAGGTACCAATG 2104
1591 SerLeuSerSerLeuValGlyThrSerGlnAsnPheAsnGluGluPheLe 1607

```

|||||
2105 TCACATCTCCTCGTGGGACATCTCAGAAATTTAATGAAGAAATCTT 2154
1607 uatgagatserleuylsThrIleleuthrtyrIagiuglunaspleuGluL 1624
|||||
2155 AGACGCTTCTTAAGACATAATTCACATATGCTGAGAGAGATCTGGAAAT 2204
1624 euargluthrThrPheProaspGlnValGlnAspleuValPheasneu 1640
|||||
2205 TGAGGGAACAACATTTCTGATCAGGTCAGAGATCTGTTTTCATATC 2254
1641 HisMetIleleuserasphThrValylsMetLysGlnHisGlnAspPr 1657
|||||
2255 CATATATTCTTCTGATCTGTAAGAAATGAAGAACACACAGAGATCC 2304
1657 oglumetleuIleaspleuMetTyrArgIleAlaLysglYtyrGlnThs 1674
|||||
2305 TGAATATTGATTCATATGATGATGATGATGATGATGATGATGATGAT 2354
1674 gATProGlu...ArgLeuthrThrPleuGlnAsnMetAlaGlyshHisSer 1689
|||||
2355 CTCACAGATCTGGGATTCGACCTGTTGCAGACATGGCAGGCAAGCACA 2404
1690 GluArgSerAsnHisAlaGlnAlaGlnCysLeuValHisSerAlaAl 1706
|||||
2405 GAACGAGCAATCATGCTGACAGCTGCACAGTCTAGTCCACTCAGCAGC 2454
1706 AleuValAlaGluTyrLeuSerMetLeuGlnAspArgLysTyrLeuProV 1723
|||||
2455 ACTGTGCTGTAATATTAGCATGCTGGAGACCGGAATATCTTCCG 2504
1723 aGlyCysValThrPheGlnAsnIleSerSerAsnValLeuGluLysSer 1739
|||||
2505 TGGATGCTTAACATTCAGAAATATTCATTAATGTTTAAACAACATCT 2554
1740 AlaValSerAspAspValValSerProaspGluGluGlyIleCysSerG 1756
|||||
2555 GCGGCTCAGATGATGCTATCTCCAGATGAAGAGATATCTGCTCGG 2604
1756 yLysTyrPheThrGluSerGlyLeuValGlyLeuLeuGlnAlaAla 1773
|||||
2605 AAAATCTTACTGAGTCAAGACTCTGTGGATTACTGGAACAAGACAGCTG 2654
1773 lAserPheSerMetAlaGlyMetTyrGlnAlaValAsnGluValTyrLys 1789
|||||
2655 CTTCCTTCTCTATGCTGCTGCATGTAAGACAGTAAATGAAGTAAACAA 2704
1790 ValLeuIleProIleHisGlnAlaAsnArgAspAlaLysLysLeuSerTh 1806
|||||
2705 GTRCTTATTCCTATTCATGAGCTAATCGGATGCAAGAAACCTATCCAC 2754
1806 rIleHisGlyLysLeuGlnGlnAlaPheSerLysIleValHisGlnSerT 1823
|||||
2755 AATTCATGCTTAACCTCAAGAGCATTCAGCAAAATGTTCTATCAGAGTA 2804
1823 hGlyTyrGlnArgMetPheGlyThrTyrPheArgValGlyPheTyrGly 1839
|||||
2805 CTGGCTGGGAGCGAGTGTGGCACCTATTTCTGTTGTTTATGGA 2854
1840 ThrLysPheGlyAspLeuAspGluGlnGluPheValTyrLysGluProAl 1856
|||||
2855 ACCAGATTCGGGATTTGGATGACACAGAAATTTTTCAGAGACCCGCGC 2904
1856 aIleThrLysLeuAlaGluIleSerHisArgLeuGlnGlyPheTyrGlyG 1873
|||||
2905 AATAACCAAACTTCGAGAGATATCTCAGAGATTCGAGGATTTTACGAG 2954
1873 lUaArgPheGlyLysAspValValGluValIleLysAspSerAsnProVal 1889
|||||
2955 AAAGATTTGGAGAGATGTGTTGAATCAACAGACTCTAATCTCTGTA 3004
1890 AspLysCysLysLeuAspProAsnLysAlaTyrIleGlnIleThrTyrVa 1906
|||||

```

```

3005 GACAAGTGTAAATTAGATCCTAACAGGCAATATATTCAGATTAACCTATGT 3054
1906 lGluProTyrPheAspThrTyrGluMetLysAspArgIleThrTyrPheA 1923
|||||
3055 GGAGCCATCTTTCACACATATGATGATGAAGACAGATACACCTATTTGCG 3104
1923 sPlyAsnTyrAsnLeuAlaArgPheMetTyrCysThrProPheThrLeu 1939
|||||
3105 ACAAAATTAACATCTTCGTGATTCATGATGATGATGATGATGATGATGAT 3154
1940 AspGlyArgAlaHisGlyGluLeuHisGlnGlnPheLysArgLysThrI 1956
|||||
3155 GATGCCCGTGGCCATGGGGAATTCATGAACAAATTCAAAAGGAAGCAAT 3204
1956 eleuthrThrSerHisAlaPheProTyrIleLysThrArgValAsnValT 1973
|||||
3205 TCTGACTACGCTCATCTCTTCTTCTTATTAATAACAGGCTCAATGTCA 3254
1973 hrHisLysGluGluIleIleleuthrProIleGluValAlaIleGluasp 1989
|||||
3255 CTCATTAAGAAAGAGATCATCTTAACACCAATGGAAGTTGCTATTTAGAG 3304
1990 MetGlnLysLysThrGlnGluLeuAlaPheAlaThrHisGlnAspProAl 2006
|||||
3305 ATGCAGAAAAAGACACAGAGATGGCATTTGCACACATCAGATCCCGC 3354
2006 aAspProLysMetLeuGlnMetValLeuGlnGlySerValGlyThrTyrV 2023
|||||
3355 AGACCCCAAAATGCTTCAGATGCTACTCCAGGATCTGTAGGCAACCCAG 3404
2023 aLaAsnGlnGlyProLeuGluValAlaGlnValPheLeuSerGluIlePro 2039
|||||
3405 TGAATCAGGGGCTTTGGAAGTTCGCCAGGTTTTCGTCTGAATAATACCT 3454
2040 SerAspProLysLeuPheArgHisHisAsnLysLeuArgLeuCysPheL 2056
|||||
3455 AGTAGCCCAAGCTCTTCAGACATCATTAATTAACCTGGACTCTGCTTTAA 3504
2056 sAspPheThrLysArgCysGlnAspAlaLeuArgLysAsnLysSerLeuI 2073
|||||
3505 AGATTTTACTTAAGAAGCTGTGAAGATGCCTTTAAGCAAAAATTAAGCTTAA 3554
2073 lGlyProValGlnLysGluTyrGlnArgGluLeuGlyLysLeuSerSer 2089
|||||
3555 TTGGCCGCTTCAAAAGAGATCAAGGAATTTGGGGAATATATCTTCG 3604
2090 Pro 2090
|||||
3605 CCT 3607

seq_name: gb_pat:AX172932
seq_documentation_block:
LOCUS AX172932 4144 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 59 from Patent WO0142297.
ACCESSION AX172932
VERSION AX172932.1 GI:14597917
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 4144)
AUTHORS Lu P., Garman, J.D. and Candia, A.F.
TITLE Clasp-3 transmembrane protein
JOURNAL Patent: WO 0142297-A 59 14-JUN-2001;
Arbor Vita Corporation (US)
FEATURES
source
1..4144
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="human Clasp-3 cDNA"
197..3610
CDS

```

```
/note="human CLASP-3"
/codon_start=1
/protein_id="CAC43728.1"
/db_xref="GI:14597918"
/translation="MDRSCNRMSSTHETSLDTLTGRLEPTKKLFHEELALQWVCSG
SVRESALQQAAMFEEELMVKSMVHLYFNDKLAPRKSRPERMDIALVSTIASDI
VSRRQKDEWERYLNTSLAEFLNDLSDYMDRGVPSLISCYKOVSKSLYSPNPSVL
VSLRDLPLRICSHENYVTLNPSLTPSPASPSVSSATSSQSFSTNVQOKIAN
MPELSVPEFROOHLAGIYLIELAVIILDPDAEGFLGHKKYINMHNLSHSDPRS
DPOIKAVANALYPLIGLIMETVPOLYDPTFENNOGRPCITATDYSESSESMISOT
VAMIACTSVPLTRPGSFLISTSGRHHTESSRSLLILWLKKNADTTLQK
MPTDLVSLQNLRLDLILYLCSEFEYKGRKVRFRMNSLFEKKSKDMRALEELLSI
GARQEWVRSRQGLERSPSGSAGSQENLKRMDMHWNRNTEKLSRAEIEHALI
DGNLAEPANLIIIDTLEIVYQVTSVESKESIGGLYKVLHSMACNQSAYVLOHCF
TORALVSKPEELLFEETEEQCADLRLRHGSSSIGTIRSHPSALYLIMRONEIG
NPNARYKMOYPMLSLIVGTSQMPNEBELPRSTKITIYAEPEYEDTEMKDRITFPKI
VNLHMLISTVYKKEHEDPEMLIDLMRIAKGYOTSPDLRLTTLQNNAKRSESN
HAELAOCLVSAALVAYELISLMDRKYLPVGCVFQNISSNVIEESAVSDVSPDEE
GICSGKFTESGLVGLLEQAAASFNAGMEAVANEYKVLPIHEANROAKKLSIHG
KLOEAFSKIYHOSGWMERFQYRVGFTGKFDDEQFKEPALITELKLAISHRL
EGFYGERGEDVVEVLIKDSNPDKCLDPKAVIOLTYVEPYEDTEMKDRITFPKI
YNLRREYCTPPTLDGRAGHELHBOEKRKTLITTSAPFVYIKTRRVVTHKEETILPI
EVALDEOKKTQELAFATHODPADPKMLQWLVDSGVTYNOGPLEVAOYFSLSEIISD
PKLFRHNKRLRLCPKDETYKRCEDALRNKSLIGPVQKEYORELKLSP"
```

BASE COUNT 1252 a 829 c 890 g 1173 t
ORIGIN

alignment_scores:
Quality: 6157.50. Length: 1201
Ratio: 5.131 Gaps: 1
Percent Similarity: 99.917 Percent Identity: 99.750

alignment_block:
US-09-737-246-2 x AX172932 ..

Align seg 1/1 to: AX172932 from: 1 to: 4144

```
891 LeuAsnArgSerArgSerLeuSerAsnSerAsnProAspIleSerGlyTh 907
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
5  GTAAATCGTCTTCGAAGCCTTAGTAATACCAATCCAGATATATCTGGGAC 54
907 rProthrSerProAspAspGluValArgSerIleIleGlySerLysGlyL 924
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
55  TCCACAGTCACCGAGATGAGAGTTCGATCAATCATCGGAGTAAGGCTT 104
924 euAspArgSerAsnSerTrpValAsnThrGlyGlyProLysAlaAlaPro 940
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
105  TAGATGCTCCAAATTCCTGGGTTAACACGTGGTGTCCAAAAGCTGCCCA 154
941 TrpGlySerAsnProSerProSerIleGluSerThrGlnAlaMetAspAr 957
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
155  TGGGGATCCAAACCCAGTCCAGATCCAGATCAACACAGGCTATGATG 204
957 gSerCysAsnArgMetSerSerHisThrGluThrSerSerPheLeuGlnT 974
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
205  AAGTTGAATCGTATGCTTCGCACACAGAGCTCAAGTTCTTCACAA 254
974 hIeuthrGlyArgLeuProThrLysLysLeuPheHisGluGluLeuAla 990
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
255  CATTAACGGGACGCTTACCAAAAGCTTTTTCACGAGGCGCTGGCT 304
991 LeuGlnTrpValValCysSerGlySerValArgGluSerAlaLeuGlnI 1007
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
305  TTGCACTGGGCTTTTTCGACAGTGCAGCTTCGGGAATCACCTTTGCAACA 354
1007 nAlaTrpPhePhePheGluLeuMetValLysSerMetValLHisHisLeuT 1024
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
355  AGCCTGCTCTTTTGAATTAATGTAAGACATGATGTCACACCATTTAT 404
1024 yTrpHisAsnAspLysLeuGluAlaProArgLysSerArgPheProGluArg 1040
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
405  ACTTTAATGATTAACCTTGAGGCTCCAAAGAAAGTCGTTTCCAGAACGT 454
```

```
1041 PheMetAspAspIleAlaAlaLeuValSerThrIleAlaSerAspIleVal 1057
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
455  TTCAATGATGATGATGATGACGCTCTTGTCACACACGATTCCTGTATATAGT 504
1057 lSerArgPheGlnLysAspThrGluMetValGluArgLeuAsnThrSerL 1074
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
505  TTCACGATTTCCAGAAAGACACAGAAATGTTGAGAGACTCAATACAAAGCC 554
1074 euAlaPhePheLeuAsnAspLeuLeuSerValMetAspArgGlyPheVal 1090
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
555  TTGCAATCTTTCTCAATGATCTGTTGCTGTATGACACAGAGATTTTGT 604
1091 PheSerLeuIleLysSerCysTyrlsGluValSerSerLysLeuTyrs 1107
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
605  TTTAGCCTTAATAAGTCTCTGTATAAACAAGTGTCTTCAAAAGCTTATCTC 654
1107 rLeuProAsnProSerValLeuValSerLeuArgLeuAspPheLeuArgI 1124
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
655  ATTACCGAATCCAGTGTCTGGTGCTCTTGAGCGTGGATTTTCTACGAA 704
1124 lIleLysSerHisGluHisThrValThrLeuAsnLeuProCysSerLeu 1140
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
705  TCATCTGAGTCATGAGCAGCATGTATACATTAACTTACCTGCAGCTTA 754
1141 LeuThrProProAlaSerProSerProSerValSerSerAlaThrSerG 1157
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
755  CTTACTCCACCTGCATCTCCATCACTCTTGCTTCTCTGACACATCTCA 804
1157 nSerSerGlyPheSerThrAsnValGlnAspGlnLysIleAlaAsnMetP 1174
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
805  GAGTCTGGATTTTCTACGAATGTACAAAGACCAAAATTTGCAATATGT 854
1174 heGluLeuSerValProPheArgGlnGlnHisTyrlsAlaGlyLeuVal 1190
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
855  TTGAATTAATCCGTGCTTTCCGCCACAGCATTTATTTGGCAGGACTTG 904
1191 LeuThrGluLeuAlaValIleLeuAspProAspAlaGluGlyLeuPheG 1207
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
905  TTAACAGAGCTGGCGTGCATTTTAGACCTGATGCTGAAGAGCTGTTGG 954
1207 yLeuHisLysLysValIleAsnMetValHisAsnLeuLeuSerSerHis 1224
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
955  ATTGCATTAAGAAAGCATCAATATGATGTAACAAATTTCTCTCCAGTCAG 1004
1224 spSerAspProArgTyrlsSerAspProGlnIleLysAlaArgValAlaMet 1240
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
1005  ACTCAGACCCCGGACTCTGACCTCAGATTAAGGCTCGAGTGCCATG 1054
1241 LeuTyrlsLeuProLeuIleGlyIleIleMetGluThrValProGlnLeu 1257
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
1055  TTGATCTACTCTGATGATGATATATCATGAAACTGTACTGACGCTGA 1104
1257 rAspPheThrGluThrHisAsnGlnArgGlyArgProIleCysIleAlaT 1274
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
1105  TGATTTTACAGAAACTCAATCAACAGAGCAACCAATTTGTATACCA 1154
1105 tGATTTTACAGAAACTCAATCAACAGAGCAACCAATTTGTATACCA 1154
1274 hAspAspTyrlsGluSerGluSerGlySerMetIleSerGlnThrValAla 1290
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
1155  CTGATGATTAAGAAAGTGAAGCGGAAATGATGATGAGCAACCGCTGG 1204
1291 MetAlaIleAlaGlyThrSerValProGlnLeuThrArgProGlySerPh 1307
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
1205  ATGGCAATCGCAGGAGACATCGGCTCCCTCAATTAACAAGGCTGCGAGT 1254
1307 eLeuLeuThrSerThrSerGlyArgGlnHisThrThrPheSerAlaGlu 1324
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
1255  CCTCTCAGCTCAAGAGTGGCAGGCAACACACTACTTTTCACACAGAT 1304
1324 eSerArgSerLeuLeuIleCysLeuLeuTrpValLeuLysAsnAlaAsp 1340
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
1305  CAACTGCAAGCCTTTTGAATCTGTCTACTTTGGGTTCTCAAAAATGCAGAT 1354
```

```

1341 GluThrValLeuGlnIuYsTrpPheThrAspLeuSerValLeuGlnLeuAs 1357
1355 GAACAGCTTCTACAGAAAGTGTTTACAGATCTCTCAGTCTGACGCTAAA 1404
1357 natgLeuLeuAspLeuLeuTyrLeuGlnCysValSerCysPheGluTyrLysG 1374
1405 CCGGCTATTAGATCTGCTTATCTCTGTGTCTCTCTTGTAGATATAAAG 1454
1374 LysLysValPheGluArgMetAsnSerLeuThrPheLysLysSerLys 1390
1455 GGAAGAAAGCTTTGAAACGATCAATAGCTTGACCTTTAAAGAAATCAAAA 1504
1391 AspMetArgAlaLysLeuGlnGlnAlaIleLeuGlnSerIleGlyAlaAr 1407
1505 GACATGAGACAAAGCTTGAAAGACTATTCTTGAGGACATAGTGCACG 1554
1407 gGlnGlnMetValArgArgSerArgGlnLeuGlnLeuArgSerProSerG 1424
1555 GCAGAAATGGTACGGCGAAGCCGAGACAGCTCGAGAGAGCCCATCTG 1604
1424 LysSerAlaPheGlnSerGlnGlnLeuAsnLeuArgTyrArgLysAspMetThr 1440
1605 GAAGTGCCTTTGAAAGTCAAGAAATTTGAGGTGAGAGAAAGATATGACT 1654
1441 HistPrArgGlnAsnThrGlnLysLeuAspLysSerArgAlaGlnIleG 1457
1655 CACTGGCTCAAAACACAGAGAAAGCTTGACAAATCAAGACAGAGATTCA 1704
1457 uHISGlnAlaLeuIleAspGlyAsnLeuAlaThrGlnAlaAsnLeuIle 1474
1705 ACACGAGAGACGTGATTGATGGAACCTGGCTACAGAAACCAACTAATCA 1754
1474 LeuAspThrLeuGlnIleValIleValGlnThrValSerValThrGlnUser 1490
1755 TTTTAAATACATTAGAGATTGTTGTTCAACACCTTTCTTAACGGAATCC 1804
1491 LysGlnUserIleLeuGlnGlyValLeuLysValLeuLeuHisSerMetAl 1507
1805 AAAGAGACGATCTTGTTGGAGTGTAAAGTCTACACAGCATGCGC 1854
1507 acysAsnGlnSerAlaValTyrLeuGlnHisCysPheAlaThrGlnArgA 1524
1855 CTGTAAACCAAGGCGATTATCTCAACACAGTGTTCGTACACAGAGAG 1904
1524 AlaLeuValSerLysPheProGlnLeuLeuPheGlnGlnIuThrGlnGln 1540
1905 CCTGGTTCAAAAGTTCTCTGAACCTTATTTGAAGAAAGACAAAGCAG 1954
1541 CysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSerIleG 1557
1955 TGTGCTGATTATGCTCTCAGGCTTCTCCGACACTGTAGCAGTAGCATCG 2004
1557 yThrIleArgSerHisProSerAlaSerLeuTyrLeuLeuMetAlaGlnA 1574
2005 TMCATACGGGTACACCCACAGTCCCTTACCTACTAATGAGGACAA 2054
1574 snPheGlnIleGlnAsnAsnPheAlaArgValLysMetGlnValProMet 1590
2055 ACTTTCAGATTGGGAATACTTGGCAGGGTTAAATGACAGGTACCATG 2104
1591 SerLeuSerSerLeuValGlyThrSerGlnAsnPheAsnGlnIuPheLe 1607
2105 TCACATATCTCTGTTGGGACATCTCAGAAATTTAATGAAGATTTCTT 2154
1607 uArgArgSerLeuLysThrIleLeuThrTyrAlaGlnGlnAspLeuGln 1624
2155 AAGAGCTTCTTAAGACTATATGACATATGCTGAAGAAGATCTGGGAA 2204
1624 euArgGlnThrThrPheProAspGlnValGlnAspLeuValPheAsnLeu 1640
2205 TGAGGGAACAACATTTCTCTGATCAGGTCCAGGATCTGTTTCAATCTC 2254
1641 HisMetIleLeuSerAspThrValLysMetLysGlnHisGlnGlnAspPr 1657

```

```

2255 CATATGATCTTTCTGATACTGTGAATGAAAGACCCAGGAGATCC 2304
1657 oGlnMetLeuIleAspLeuMetTyrArgIleAlaLysGlnTyrGlnThrS 1674
2305 TGAATGTGATTGATCTAATGTACAGAAATGGCAAGGCTTACAGACCT 2354
1674 ePrProGlu...ArgLeuThrTrpLeuGlnAsnMetAlaGlyLysHisSer 1689
2355 CTCGAGATCTCGATTGACCTGGTTGCAGAACATGGCAGGCAACCTCA 2404
1690 GluArgSerAsnHisAlaGlnAlaAlaGlnCysLeuValHisSerAlaAl 1706
2405 GAACAAACATTCATCGCTGAAGCTGCACAGCTGTAGTCCCTCAGACAG 2454
1706 AlaValAlaGlnTyrLeuSerMetLeuGlnAspArgLysTyrLeuProV 1723
2455 ACTTGTCTGAATATTTGACATGCTGAGAGACCGGAATATCTTCTCTG 2504
1723 aGlyCysValThrPheGlnAsnIleSerSerAsnValLeuGlnGlnUser 1739
2505 TGGATGTGTAAACATTTTCAGAAATTTTCATCTAATGTTTAAAGAAATCT 2554
1740 AlaValSerAspAspValIleSerProAspGlnGlnGlyIleCysSerG 1756
2555 GCGGCTCAGATGATGTGTGTTCTCCAGATGAAGAGATTCGTCTGTGG 2804
1756 LysTyrPheThrGlnSerGlyLeuValGlyLeuLeuGlnAlaAlaAla 1773
2605 AAAATACTTACTAGTCAGACACTTGTGGATTACTGGAACAAGACACTG 2854
1773 LasePheSerMetAlaGlyMetTyrGlnAlaValAsnGlnValTyrLys 1789
2655 CTCTCTCTCTATGCTGGCTGTGATGAGCAGTAAATGAAGTTTACAAA 2704
1790 ValLeuIleProIleHisGlnAlaAsnArgAspAlaLysLysLeuSerTh 1806
2705 GTACTTATTCCTATTTCATGAAGCTAATCGGGATGCCAAAGAACTATCC 2754
1806 rIleHisGlnLysLeuGlnGlnAlaPheSerLysIleValHisGlnSert 1823
2755 AATTCATAGTAACTTCAGAGACATTCACGAAATGTTTCATCAGAGTA 2804
1823 hrgLysTrpGlnArgMetPheGlyThrTyrPheArgValGlyPheTyrGly 1839
2805 CTGGCTGGAGCGGATGTTGGCACCTATTTCGTGTTGTTTATGGA 2854
1840 ThrLysPheGlnAspLeuAspGlnGlnIuPheValTyrLysGlnProAl 1856
2855 ACCAAGTTCCGGGATTTGATGAACAAGAAATTTGTTACAAAGGACCTGC 2904
1856 aIleThrLysLeuAlaGlnIleSerHisArgLeuGlnGlyPheTyrGly 1873
2905 AATTAACCAAACTTGACAGATATCTCACAGATTGGAGGATTTTACGAG 2954
1873 LuArgPheGlnGlnAspValValGlnValIleLysAspSerAsnProVal 1889
2955 AAAGATTGGAGAGATGTGTTGAATCAAAAGCTTAATCCCTGA 3004
1890 AspLysCysLysLeuAspProAsnLysAlaTyrIleGlnIleThrTyr 1906
3005 GACAAGTGTAATATAGATCTTAACAAGGCAATATATTCAGATACGATG 3054
3055 GGAACCACTACTTGCACATATGAGATGAAGAGACAGATCACCTATTTGG 3104
1906 LgluProTyrPheAspThrTyrGlnMetCysAspArgIleThrTyrPheA 1923
1923 sPLysAsnTyrAsnLeuArgArgPheMetTyrCysThrProPheThrLeu 1939
3105 ACAAAATTTCAATCTTCGTCGATTCATGTACTGTACACCCCTTACTTTA 3154
1940 AspGlyArgAlaHisGlnGlnLeuHisGlnGlnPheLysArgLysThrI 1956

```



```

162 IuAlaProAspGlyAsnSerTyrGlnAspSpGlnAspAspLeuLysArg 178
    |||
407 CTTGGAG.....TCAGTGAACCCGCTGCTCAGGCAGGCCCCGCCAC 450
    |||
179 ArgSerMetSerIleAspSerThrProAlaGlySerThrAlaCysSerI 195
    |||
451 TTAACGCTGCTGCGAGCTGTCTGGAAAGCCCGCTCAGTCCCTGTGA 500
    |||
195 ePheAspLeuLysAsnSerLeuProAspAlaLeuLeuProAsnLeuLeu 212
    |||
501 CTTGACCTCCGACGCTCAGCCTACACAGCGGCTAGAAACCTCTCTCC 550
    |||
212 spArgThrProAsnGlnGluIleAspArgGlnAsnAspAspGlnArgLys 228
    |||
551 AGCAAGTGAAGTCCGAGGAGCTTGAAGACAGAAAGAGAGGCCCGCAG 600
    |||
229 SerAsnArgHisLysGluLeuPheAlaLeuHisProSerProAspGlu 245
    |||
601 ACCAAGAGGAGGCGAGCTCTTGGCCCTTACCCATCAGTGCAGCAGGA 650
    |||
245 uGluProIleGluArgLeuSerValProAspIleProLysGluHisPhe 262
    |||
651 GAGTGTGTGGAATATACGTCACGTACCAAGATGTCCCAAGAACCTCG 700
    |||
262 LysIleArgLeuLeuValLysCysLeuSerLeuLysPheGluIleGlu 278
    |||
701 GCACAGAAATATGTCAAGTCTGACCTTGAAGTTGAGATTGAAT 750
    |||
279 GluProIlePheAlaSerLeuAlaLeuTyrAspValLysGluLysLys 295
    |||
751 GAGCCCTGTTGGCAGCATGTGCCCTCTACGATGTTAAAGAAAGGAAAA 800
    |||
295 sIleSerGluAsnPheTyrPheAspLeuAsnSerGluIleMetLysGlu 312
    |||
801 GATCTGAGAAATTTTCACTGCTGACCTGAACTGTGCACGTTCAAAAG 850
    |||
312 euleuArgProHisValProProAlaAlaIleThrThrLeuAlaArgSer 328
    |||
851 TTTCTGCGAGTCCACAGCCTTCTGAGCCGACATCAAGTCAGCGAGAT 900
    |||
329 AlaIlePheSerIleThrTyrProSerGlnAspValPheLeuValIle 345
    |||
901 GCGGTCTTCACTCAGTACCTACCCGCTCTCAGACATCTACCTGTGTCA 950
    |||
345 sLeuGluLysValLeuGlnGlnGlyAspIleGlyLysGluLysAlaGlu 362
    |||
951 GATTGAAAAGTCTCTCAGCAGGAGATATGTGAGACTGTGCAGACCT 1000
    |||
362 yrMetIlePheLysGluAlaAspAlaThrLysAsnLysGluLysLeu 378
    |||
1001 ACACGGTATCAAAAGAAAGTGAATGTGGAAGAGTAAGAAAGATTGAA 1050
    |||
379 LysLeuLysSerGlnAlaAspGlnPheCysGlnArgLeuGlyLysTyr 395
    |||
1051 AAACVTAACCTCCAAAGCTGAATCTTGTCCAGCGTTGGGGAATACCG 1100
    |||
395 gMetProPheAlaThrAlaIleHisLeuMetAsnIleValSerSer 412
    |||
1101 GATGCCCTTGGCTGGGACCCATTAAGCTTATCAAGCTTCTTC...AATG 1147
    |||
412 laGlySerLeuGluArgAspSerThrGluValGluIleSerThrGly 428
    |||
1148 TCTCCACCTTGAAGAGGAGTAAGTGAATGTGACTGTGTGTGG... 1194
    |||
429 ArgLysGlySerThrSerGluArgArgAsnSerSerIleValGlyArg 445
    |||
1195 ...AGAACCCAGTGGGTGAAGGAGGAGCATTTGCC...CAATCTAAGAG 1238
    |||
445 gSerLeuGluArgThrThrSerGlyAspAspAlaCysAsnLeuThrSer 462
    |||
1239 GCTTTTGAAAAGAGCCCTCTCTTGGAGGAAATAGGGGTGATCAACT 1288
    |||
462 heArgProAlaThrIleuThrValThrAsnPheLysGlnGluLys 478

```

```

1289 TCAAAACCTCCACCTGTAGCGCTTAGACCTTTTCAAGCAGAGAGAGAT 1338
    |||
479 ArgLeuSerAspGluAspLeuTyrLysPheLeuAlaAspMetArgArg 495
    |||
1339 CGCCTTACCGATGAAGACTTATCAAGTTTATCTACCTACTCAAAAGATC 1388
    |||
495 oSerSerValLeuArgArgLeuArgProIleThrAlaGlnLeuLysIle 512
    |||
1389 ATCATCTTACAGAGACAGTCAAGTCAATTCAGGCTGTGTAAGACTGG 1438
    |||
512 spIleSerProAlaProGluAsnProHisTyrCysLeuThrProGluLeu 528
    |||
1439 AGATTCTTACAGCTCCAGAGATCAATTCATGCTGTGACTCCGAAATG 1488
    |||
529 LeuGlnValLysLeuTyrProAspSerArgValArgProThrArgGlu 545
    |||
1489 CTGCCCGGAAACCTTCTGAAAACCGGACACGCGGACCAAGAGAT 1538
    |||
545 eLeuGluPheProAlaArgAspValTyrValProAsnThrThrTyrArg 562
    |||
1539 TTTGGAATTTCAACACAGAGATATATGTCCTCACACTGTGTACGAA 1588
    |||
562 snLeuLeuTyrIleTyrProGlnSerLeuAsnPheAlaAsnArgGln 578
    |||
1589 ACCTTCTATGCTACCCACAGAGCGCTGAACCTTGAACAAACTGACA 1638
    |||
579 SerAlaArgAsnIleThrValLysValGlnPheMetTyrGlyLysArg 595
    |||
1639 TCAGCCCGGAACATTTACAAATTAAGATCAGTTTATGTGTGAGAGAT 1688
    |||
595 oSerAsnAlaMetProValIlePheGlyLysSerSerCysSerGluPhe 612
    |||
1689 TAGCAATCGATGCCGCTGATCTTTGAAATTCAGGCGGCTGAATTC 1738
    |||
612 eLysGluAlaTyrThrAlaValAlaTyrHisAsnArgSerProAspPhe 628
    |||
1739 TGCAGGAAGTGTACACAGCTGTATCATACATATATAGTCTCCGACTT 1788
    |||
629 HisGluIleLysValLysLeuProAlaThrLeuThrAspHisHis 645
    |||
1789 TATGAGAAAGAAATTAAGCTCCCGCTAAGCTCAGTAATATACCA 1838
    |||
645 sLeuLeuPheThrPheTyrHisValSerCysGlnGlnLysGlnAsnThr 662
    |||
1839 CTTCTGTTCACCTTACCATATACACTGTACACAAAGCAAGAGACCT 1888
    |||
662 rLeuGluThrProValGlyTyrThrThrIleProMetLeuGlnAsn 678
    |||
1889 CCGTGAAGAACTCTCGGATATTCATGAGCTGCCAATTCCTTAATGAA 1938
    |||
679 ArgLeuLysThrGlnPheCysLeuProValSerLeuGluLysPro 695
    |||
1939 CGTCTTCAAACTGGATCTTACTGTCTCCAGTTGCTTGGAAAAATTGCC 1988
    |||
695 oGlnAlaTyrSerValLeuSerProGlu...ValProLeu..... 709
    |||
1989 ACCCACTACTCCATGCAATTCCTGCTGAGAAAGTCCATTACAGAACTC 2038
    |||
709 LysMetLysThrValAspAsnHisLysGlyValPheAsnValGluVal 725
    |||
2039 CCATTAAAGTGGCGTGAAGAGCATTAAGGAGATTATTAATTGAAGTCCA 2088
    |||
726 AlaValSerSerIleHisThrGlnAspProTyrLeuAspLysPhePhe 742
    |||
2089 GCTGTCTTCTTGTACACACCAGGACCAACACCTGTGGAAGTCTTCAC 2138
    |||
742 aLeuValAsnAlaLeuAspGluHisLeu...PheProValArgIleGly 758
    |||
2139 CTTGTGCACCTCCCTGGAGAGCCAGGTGACCTTCCCATCCTGGCTCG 2188
    |||
758 spMetArgIleMetGluAsnAsnLeuGluAsnLysLeuLysSerSerIle 774
    |||

```

2189 ATCAGAAAATCAGCAGATGGCGCTGGAGCATGAGCTGAACCTCAGCATC 2238
775 SerAlaLeuAsnSerSerGlnLeuGluProValValArgPheLeuHisLe 791
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2239 ATCTGCTGAAGTCTCCCGCTGGAGCGGCTGCTCTTCGCGACCT 2288
791 uLeuLeuAspLysLeuIleLeuLeuValIleArgProProValIleAla 808
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2289 GGTGTGTCAGAGCTCTTCAGAGCTGCTCCGTCAGCCCATGTGTCATGCTG 2338
808 lYgLIleValAsnLeuGlnLeuGlnIleAspPheGluAlaMetAlaSerLe 824
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2339 GCCAGACAGCCCACTTCTCCCATGTTGCTTGAGTCCGTGCTGAGCCATC 2388
825 lLeAsnArgLeuHisLysAsnLeuGluLysAsnHisAspGlnHisGlyAr 841
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2389 GCCAACAGCTGACACACAGCAGAGCCTGAGCAAGACGACATGGGAG 2438
841 gAsnSerLeuLeuAlaSerTyrIleHisTyrValPheArgLeuProAsnT 858
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2439 GAACGTGCTGTGCTCTCTACGTGACCTACGCTCTCCGCTGCGCAGAGG 2488
858 hTyrProAsnSerSerSerProGly...ProGlyGlyLeuGlyGlySer 873
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2489 TGCAAGAGGATGTGCCAAGTCAGGCGCTCCCACTGCCCTCTCTAGACCCT 2538
874 ValHisTyrAlaThrMetAlaArgSerAlaValArgProAlaSerLeuAs 890
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2539 CGGAGCTACACACAGTATGGCCGACATCAGCTGCTGTGAGTTCAAA 2588
890 nLeuAsnArgSerArgSerLeuSerAsnSerAsnProAspIleSerGlyT 907
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2589 GCTGCTGAGGCGCGGATGATGACAGCATACCCAGACCTCGCGGGGA 2638
907 hPProThrSerProAspAspGluValArgSerIleIleGlySerSerGly 923
||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2639 CACACTCCGCGACAGCAGCAGAGTGAAGAACATCATCTTCCAAAG... 2685
924 LeuAspArgSerAsnSerTyrValAsnThrGlyGlyProlYsAlaAlaPr 940
2685 2685
940 cTriPolySerAsnProSerProSerAlaGluSerThrGlnAlaMetAspA 957
|||:::|||||:::|||||:::|||||:::|||||:::|||||
2686 ATCGCCGATC 2695
957 rGSerCysAsnArgMetSerSerHisThrGluThrSerSerPheLeuGln 973
||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2686 GCAACTGCAGCCGAAATGCTTACTATTTGCTCTGCGCAGTAgTAgTCCCA 2745
974 Thr.....LeuThrGlyArgLeuProThrLysLysLeuPheHisGluG 988
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2746 AGTTCACCTGCAGCCCAAGGCCAGCCAGCAAAAGCATTTCCATGAGGA 2795
988 uLeuAlaLeuGlnTrrPValValLysSerGlySerValArgGluSerAlaL 1005
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2796 GCTTGCCCTTACAGATGTGTGTCAGACCGGA..... 2826
1005 euGlnGlnAlaIrrPhePhePheGluLeuMetValLysSerMetValHis 1021
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2827 ATGCTGAAGAGCATGGCCAG 2847
1022 HisLeuTyrPheAsnAspLysLeuGlnAlaProArgLysSerArgPhePr 1038
||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2848 CACGTACATTAACATGACAAAGCGAGTTTTCGAGAGCTCGTTTTC 2897
1038 cGluArgPheMetAspPheIleAlaLeuValSerThrIleLeuAspA 1055
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2898 TGACCGTTTCATGATGATPACTACTATTGTAAATGTGTGTCACCTCG 2947
1055 sPileValSer.....ArgPheGlnLysAspThrGluMetValGlu 1068
||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2948 AATTTGACGCCCTTTTACTAAACACAGAAAGAAATGAACAGCGGGA 2997

1069 ArgLeuAsnThrSerLeuAlaPhePheLeuAsnAspLeuLeuSerValMe 1085
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2998 AAGATGAACATCAGCCTGGCTTCTCTGTGATGACCTTCTGCCCTCAT 3047
1085 tAspArgGlyPheValPheSerLeuIleLysSerCysTyrLysGlnValS 1102
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3048 GGATCGGGGGCTTTGTGTTTAACCTCATCAGACATTTTCACGACGCTGT 3097
1102 eSerLysLeuTyrSerLeuProAsnProSerValLeuValSerLeuArg 1118
||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3098 CAGCCAGCTCACTAATCTTCCA.....AGCTCATTTTCATGAGG 3138
1119 LeuAspPheLeuArgIleLecSerHisGluHisTyrValThrLeuAs 1135
||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3139 CTAGAGTCTCTGAGAAATCTCTGTAGCATGACATTTACTCATCTGAA 3188
1135 nLeu...ProCysSerLeuLeuThrProProAlaSerProSerProServ 1151
|||:::|||||:::|||||:::|||||:::|||||:::|||||
3189 CTTTTTTTTTATGAATGCTGATAGTCTCTCCAAACATCTCTGTCTTCCA 3238
1151 aSerSerAlaThrSerGlnSerSerGlyPheSerThrAsnValGlnAsp 1167
||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3239 TATCT.....TCCAGAACTCAAGCTCTGCTCAGCTTCCAGGAC 3279
1168 GlnLysIleAlaAsnMetPheGluLeuSerValProPheArgGlnGlnH 1184
||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3280 CAGAGATGCCCGACGATGTCGATGCTCGAGTACCGCAGCAGAGA 3329
1184 sTyrLeuAlaGlyLeuValLeuThrGluLeuAlaValIleLeuAspProA 1201
||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3330 CTTCCTCACCGGGCTCTCTTTCACAGAACTGCTGCTGCCGTGAGTCCG 3379
1201 sPAlaGluGlyLeuPheGlyLeuHisLysLysValIleAsnMetValHis 1217
||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3380 AAGGGAGAGAAATCAGCAAAAGTCAAAAGAAAGCTCTGAGTCAATTCAC 3429
1218 AsnLeuLeuSerSerHisAspSerAspProArgTyrSerAspProGlnI 1234
||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3430 AGCTGTGTAAGTCTCAGCAGCTGAGCCGACCGCTGTGTCAAACAGAGT 3479
1234 eLysAlaArgValAlaMetLeuTyrLeuProLeuIleGlyIleIleMetG 1251
||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3480 GAAGGTCAAAATCGCCGCTTACCTACCTAGTTGATGTCATCATTTTGG 3529
1251 luthrValProGlnLeuTyrAspPheThrGluThrHisAsnGlnArgGly 1267
||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3530 ATGCTTTCACAGACTCTGTGACTTTTACAGTTGACAGATCTCGAGATAC 3579
1268 ArgProIleCysIleAlaThrAspAspTyrGluSerGluSerGlySerMe 1284
||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3580 CGC.....ACAGTGGCTCGATGAAAGAAAGAGGAGCGGCTGC 3620
1284 tIleSerGlnThrValAlaMetAlaIleAlaGlyThrSerValProGlnL 1301
||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3621 CATTAAACCAAAATGTGCTGTGCGCATAGCAGGGAATTAATTCATTTGA 3670
1301 euthrArgProGlySerPheLeuLeuThrSerThrSerGlyArgGlnHis 1317
||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3671 AAACA.....AGTGAATATGTCTGTCTTCTGCTCATTAAGCAGTAC 3714
1318 ThrThrPheSerAlaGluSerSerArgSerLeuLeuIleCysLeuLeuTr 1334
||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3715 AACATGCTGAACGCGGACACTACTCGCAACCTCATGATCTGCTCTCG 3764
1334 pValLeuLysAsnAlaAspGluThrValLeuGlnLysTrpPheThrAspL 1351
||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3765 GATCATGAAAATGTGATCAGAGCCTCATTTAGAGATGATTTGCGACG 3814
1351 euSerValLeuGlnLeuAsnArgLeuLeuAspLeuLeuTyrLeuCysVal 1367
||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3815 TGCCATCAACGCGACTCAACAGAGATTTTATGATCTACTTTTCATCTGTGC 3864


```

5641 ATCAGACAGCAGATCAGCGTCATCCAGAAAGAGAGTTGTTTACACC 5690
1982 oilegluvalalalegluaspmetglnlyslslythrnglnleualap 1999
|||||.....:|||||:|||||
5691 GATTGAAGTGGCATTTGAAGACATGAAAGAGAACACCTGCAGTTAGCG 5740
1999 healatHhISgluasprolaaaprpolyrsmelenglmethylleu 2015
|||:.....||| ||| |||||||||.....|||
5741 TTGGCATTAAACGAGGAGCGCTGATGCAAAAGATCTTCAGATGGTGTG 5790
2016 glnlgyservaIglythrValasnnglnlyProleugluvalalagl 2032
|||||.....:|||||:|||||
5791 CAAGGCTCTGTGGAGCTACTGTTAAATCAGGACCACTGGAAGTACCCA 5840
2032 nValpheuSerIuIleProSerAspPolyrsmleupheargHhISA 2049
|||||.....:|||||:|||||
5841 AGTGTTTTGGCTGAATTCCTGTCGATCCAAAACTCTATCGACATWCA 5890
2049 sHlyleuArgleucyspHelysAspHethlylArgcysluaspala 2065
|||||.....:|||||:|||||
5891 ACAGATTGAGGTTATGCTTTAAGGAATTCATCATGATGATGTTGTAAGCT 5940
2066 leuArglyAsnlySerleuIleglyProvalglnlysglutyrglnar 2082
5941 GTAGGAGAAAACAGCGCTCTCATCAGCGCAGCACGAGGGAATTCAGCA 5990
2082 gSglueuglylys 2086
5991 GGAACCTCAAAAAG 6003
seq_name: gb_pr:AB037816
seq_documentation_block:
LOCUS AB037816 4886 bp mRNA linear PRI 14-MAR-2000
DEFINITION Homo sapiens mRNA for KIAA1395 protein, partial cds.
ACCESSION AB037816
VERSION AB037816.1 GI:7243170
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Nagase,T., Kikuno,R., Ishikawa,K.I., Hirosewa,M. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes.
XV. The complete sequences of 150 new cDNA clones from brain which
code for large proteins in vitro
DNA Res. 7 (1), 65-73 (2000)
JOURNAL
MEDLINE 2 (bases 1 to 4886)
REFERENCE Ohara,O., Nagase,T. and Kikuno,R.
AUTHORS Direct Submission
TITLE Submitted (31-JUN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdna@ofekazusa.or.jp,
URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)
FEATURES
Source
1..4886
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="hj07927"
/tissue="brain"
/clone_lib="pluescriptII SK plus"
1..4884
/gene="KIAA1395"
<1..>4884
/gene="KIAA1395"
/note="Start codon is not identified."
/codon_start=1
/product="KIAA1395 protein"
/protein_id="BAA92633.1"

```

```

/db_xref="GI.7243171"
/translation="DARTMAASERAAHKNIRVAAEVKQVRENSGSHRRRS
SSGLQPLEVEEPLDEPDLISRPDAPCGRLQVLEPADDLELQIPRCRTTEG
IPKOEKLDQVRAAVEMYIDBWIVHRYGYLSAAYSPVTDIQREKQGLPQVQ
DASDEKEDGVEDSNDSSRRSGSPEDTPRSSGASSIPDLRNLAOSLPSLIERAAPD
VDRENTLRROHREPALLTYLPADDEAAVERCSRPRPREHFGORILVCLSKFEL
EIEPIGILALVDYREKKKTSENFYPLNDSMKGLLRAGTHPAISTLRASITFVY
YPSDFITVLTKLEVLDQGGDISECCETWMLKVDYPAKKNEKLEKLAAQECTRLG
RYRMPFAMTAVHLNANIVSSAGQUDRDSDBERNPAWTDRRRQDRASSGDACS
SGEPFALTVTFNFKOAERLSDEKPLDAERNPSSLRLRLPVTALQKIDISPA
ENPFICSPELHLHKPYDPGRGPTKEVLEPAEVAPPTYSNLLYVPHSNFESS
ROGVRLNLAVOYMGEDPSOALPYIEGSSCEPFRLEAFTPVYHNKSPFEERK
LHPACYTEHNHLLFTFYHVSCORPCRTLETPYGFMTPLDQHRRLRTGFCPLVSV
DQPPPSVULPDUVALQMKRVGDHKVEFYSVELTAVSVNHODPYLDKFPVLVYLEE
GAFPERLKDVIVLSEGNVEQELRSLAALRLASPEPLVAFSHHVDKLVRLTRPITL
GOIYNLRGAFEMAAHVSVLHRSLEAODARGCPOLAAVHVAFLRPGTEPGLPOG
APPTVOAATLARGSGRPASLYLARSKSISSSNDLVAPEVSDVESRIILASRLHE
EIALOWYSSAVREAILLOHAFMEPLMVKSMALHLLQGGDIPRLKRFGRDLDDI
TALVSGVLEYITVPHKDVLEAEFLNMAFSLDILSYDGRGVEFLVRAHYKQAVN
RLQSSPMPAALLITRMETRLTCSHENYVTLNDCPLSPASPSPEVSTTSOSSTFR
SSQAPDKRVTSMFELSGEPFQOHFLAGLITLLEALALEPEEAGFLHKKATISVHSD
LCGHDTPREYAEATVAKRVALELYPLLSIADTLPLELAEPEEGGORSRLASMDSPD
EGEDIDAGTINPSVAMAIAAGPLAGSASISQPPRLASRGALSSESSRTLLACV
WYLNTERPITALORATQTLIPOLGRILDIYLICLAEPYCKKAFERNLSITFEKSLD
MKALEERILICTIGARQDMVARSREPSFGPENVVRNRKSTTHNKSIDRDKRDE
EHEALVEGNLATEASLVLDLLETIVOTVMLSEKRESYLVAVLAVLTSLSAOSAL
LOHGLATQRALVSKFELPEDEPDELADLCRLRLHOGSITSTRHASLSTYLLMK
ONFEGHNFAVKMQVMYSSLSIVGTQNFQFEELRSLKTLITYAEEDMGLRDSFPR
EVOYDILNMLHMLITDYKMKHEQEDPEMLIDLMLRYLARGQSGPDLRLTLQNMAGR
HAEIGNHAEAAQCYNVHA"
BASE COUNT      941 a      1603 c      1425 g      915 t
ORIGIN

alignment_scores:
      Quality: 5472.50      Length: 1711
      Ratio: 3.835      Gaps: 11
      Percent Similarity: 83.402      Percent Identity: 61.718

alignment_block:
US-09-737-246-2 x AB037816 ..

Align seg 1/1 to: AB037816 from: 1 to: 4886

2 Alag1ArGArGAlaPheAlaGlnIyLSerArgThValAla1a1 18
:::|||||:::|||||:::|||||:::|||||:::|||||:::
22 TCGGACGCGCGCCCTTCGGCGACAGATCAACAGAGCGGTGGCCGAGA 71
18 vAla1ArgLySGlnIleSerGlyGlnTySerGlySerProGlnleuLeul 35
|||||:::|||||:::|||||:::|||||:::|||||:::
72 GGTGGGAAACAGAGTGTCCCGGAGACGATGAGTGTCCCGCCACATCCAGCA 121
35 yAsnleuAsnIleValGlyAsnIleSerIshStIrrThrValProleu 51
:::|||||:::|||||:::|||||:::|||||:::
122 GG.....CGGTGCAGCAGGCTCCCTGGGGGCTCCACCTG 153
52 ThnGluAlaValAspProValAspLeuGluAspTyTleuIleThrHis 68
|||||:::|||||:::|||||:::|||||:::|||||:::
154 ACTGAGAGTTTCAGGCCCTTCGAGACTTGGAGAGATGTACTTTCGACCGCGCC 203
68 OlenuAlaValAspSerGlyProleuArGAspLeuIleGluPheProPro 85
|:::|||||:::|||||:::|||||:::|||||:::
204 ACCAGATGTGAGCGCCGCGCCCTTCAGAGGACGTGTAAGTATCCACACTG 253
85 sPaSPllEGluAlaValTySerProArGAspCySaTgThrLeuValSer 101
|||||:::|||||:::|||||:::|||||:::|||||:::
254 ATGACTGTGAGCTGTCTGCTCAGACCCCGGAAATGACCGAGACGAGACC 303
102 AlAValProGluSerGluMetArGAspProHisIleValArGAspCysIleAr 118
:::|||||:::|||||:::|||||:::|||||:::|||||:::
304 GGGATCCCAAGATGAAACAGTGAATGCCACAGGTGAGGGCCGCGGTGGA 353
118 gSerTyTThnGluAspTrpAlaIleValIleArgLyTyThHisLysLeuG 135

```

1254 CTCTGAC.....TCGGAGGGCGAGCGCCGGCCAGCCTTGACAG 1291
435 LuArGaRgsnSerSerIleValGIyArGaRserLeuGIuArGlyThr 451
1292 ACCGCCCCCGT.....CGGGGGCCCCAGAGACCGGGGAGT 1326
452 SerGIyAspAlaIcysAsnIleuThrSerPheArGProAlaThrLeuTh 468
1327 AGTGGGACACGCCCTGCACCTTCTGCTTCCGCTCCGACGACGCTTAC 1376
468 rValThrAsnPhehelsGIngluIyAspArGLeuSerAspGIuAspL 485
1377 TGTCACAACCTCTTTAAGCAGAGGCTGAGCAGCATGAGAGGAGACC 1426
485 eUyTrIySPheLeuAlaAspMetArGaRProSerSerValIleuArGa 501
1427 TCTTCAAGTTCCTGGGTGACATGAGCGCCGCTGCTCCGTCGCGGCA 1476
502 LeuArGProIleThrAlaGInLeuLysIleAspIleSerProAlaProGI 518
1477 CTACGTCCTGTGACTGCCCGAGCTCAAGATGCACATTTCTCCGGCTCTGA 1526
518 uAsnProHisTyRcysLeuThrProGluLeuGInValIySLeutyR 535
1527 AAATCCCACTTCTGCCTCTCCCTGAGCTGCTTCATATCAACCCCTACC 1576
535 roAspSerArGValArGProThrArGInuIleLeuGIuPheProAlaAr 551
1577 CGGACCCAGGGGCGGCCCCACCAAGAGATTCTGAGTTCGCCGCCGCC 1626
552 AspValIyValProAsnThrThyArGAsnIleuLeuTyRILeTyR 568
1627 GAAGTCAATCCCCCATACACACTACAGAACTCTGTACGCTGACACC 1676
568 oGInSerLeuAsnPheAlaAsnArGIngluIySerAlaArGAsnIleThv 585
1677 GCACAGCCTCACTTCAGCAGCCGCCAGGCTCCGCGCCACACTTGCTG 1726
585 alIySValGInPheMetTyRGIyLAspProSerAsnIleMetProVal 601
1727 TCGGAGTGCATGATGACAGCGGAGAGACCCAGCCAGGCTCTGCGGTC 1776
602 IlePheGIyLysSerSerCysSerGIuPheSerIySGluAlaTyRThrAl 618
1777 ATCTTGGCAAGTCCAGCTGCAGTGAATTACCCGGAGGCTTCACACC 1826
618 aValValTyRHisAsnArGSerProAspPheHisGIuGIuIleYsValL 635
1827 GGTGGCTTACCATTAACAAGTCCGCCAGTTCACAGGAGTTCACAGCTGC 1876
635 ySLeupProAlaThrLeuThrAspHisHisIleuLeuPheThrPheTyR 651
1877 ATCTTCAGGCTGCTGACAGAGAACATACCTGCTGTCACTTCTTAC 1926
652 HisValSerCysGIngluIySGInAsnThrProLeuGIuThProValGI 668
1927 CATGTGAGCTGCACGCCCGCGGCGGCTGACCTGCGTGAACAACCGTGGG 1976
668 yTyRThrTrPIleProMetLeuGInAsnGIyArGLeuTyRIngluP 685
1977 CTTTACTTGGATCCACTGCTGCACAGCGCCCTGAGAGACCGGCCCT 2026
685 heCysLeuProValSerLeuGIuLysProProGInAlaTyRSerValLeu 701
2027 TCTGTCTCCAGTGTGTGTGGACACGCCCGCCAGCTATTTCCGTGCTC 2076
702 SerProGIuValProLeuProGIuMetIyTyRValAspAsnHisLysGI 718
2077 ACACCGCATGTGGCGCTTCCGGGCAATGCCGTGGAGAGGTCAACAAGG 2126
718 yValPheAsnValGIuValAlaValAserSerIleHisThrGInAsp 735
2127 CGTGTCAAGTGGAGCTCACGGCCGTGTCTGTGACACCCCGAGGACC 2176

354 GATGATATATTGAGACTGGGTGTCATGTCACAGAGGATATACGTACCTGA 403
135 IyThrGIyPheAsnProAsnThrLeuAspLysGIuLysGIuArGIuLys 151
404 GTGCAGCATACAGCCCGTCACACAGACACACCGGGAGGAGACAGAG 453
152 GIyLeuProLysGIuValPheGIuSerAspGIuAlaProAspGIyAsnSe 168
454 GGCCTTCCCGCCAGGTCTTTAGCAGAGTGTCTTGAGACAGAGAGGCTC 503
168 rTyRGIuAspAspGIuAspAspLeuLysArGArGserMetSerIleAspA 185
504 CGGCCCTGAGGACTCGAAATGACTCCCGGCGTCCGCGCTCCCGGAG 553
185 sPTThrProArGIySerTrpAlaCysSerIlePheAspLeuLysAsnSer 201
554 ACACCCCTCGAAGCAGTGGTCTTACCATCTTGACCTGAGGAAACCTG 603
202 LeuProAspAlaLeuLeuProAsnLeuLeuAspArGThrProAsnGIuGI 218
604 GGAGCTGACTCATGTGCTGCTGCTGCTAGAGCGGGCGGCCCGACAGA 653
218 uIleAspArGIuAsnAspAspGIuArGIySerAsnArGHisLysGIuL 235
654 TGTGAGACGGCGCAATGAACCTTCGACGAGCAGACCGGCCCGGCC 703
235 euPheAlaLeuHisProSerProAspGIuGIuGIuProIleGIuArGIu 251
704 TGTCAACCTCTACCCGGACCTGACGAGATGAAGCCGTGAGACGCTGT 753
252 SerValProAspIleProLysGIuHisPheGIyGInArGLeuLeuValLy 268
754 ACCGCCGAGAGACCCCGGAGCATTTGGACAAAGATCTTGCTGCTAA 803
268 sCySLeuserIeLysPheGIuIleGIuIleGIuProIlePheAlaSerL 285
804 GTGTCTGTGCTCAAGTTCGAGATTGAATTGAGCCCATCTTGGGAGCT 853
285 euAlaLeuTyRAspValIySGInLysLysLysIleSerGIuAsnPheTyR 301
854 TGGCTCTGTATGATGTCGGGAGAAAAAGAGATCTCGAGAACTTCTTAC 903
302 PheAspLeuAsnSerGIuGInMetLysGIyLeuLeuArGProHisValPr 318
904 TTCCGACTGAACTCGGACTCATGAAGGGCTGTGCGGCTCATGGCAC 953
318 oProAlaAlaIleThrThrLeuAlaArGSerAlaIlePheSerIleThrT 335
954 CCACCTCCCATCTCCACCTGGCCGCTGCGCATCTTCTGTGACCT 1003
335 yRProSerGIuAspValPheLeuValIleLysLeuGIuLysValLeuGIu 351
1004 ACCCTTACACGTGACATCTTCTGCTGCTGCTCAAGTTGGAGAAAGTCTTAC 1053
352 GIuGIyAspIleGIyGIuCysAlaGIuProTyRMetIlePheLysGIuAl 368
1054 CAAGGGGACATCAGTGAAGTGTGAGCCTTACATGTTGAAGAAGT 1103
368 aAspAlaThrLysAsnLysGIuLysLeuGIuLysLeuLysSerGIuAla 385
1104 GGACACACCCAAAGCAAGAGAGCTAGAGAACTGGCGCTGGCGCCG 1153
385 sPGLuPheCysGInArGLeuGIyLysTyRArGMetProPheAlaTrpThr 401
1154 AGCAATTTGACACCGCGCTGGCGCTACCGCATGCCCTTGCTGAGACG 1203
402 AlaIleHisLeuMetAsnIleValSerSerAlaGIySerLeuGIuArGAs 418
1204 GCCGTGACCTTGGCCACATCGTAGACGCTGGGACGCTGACCGGGA 1253
418 pSerThrIuValGIuIleSerThrGIyGIuArGIyGIySerTrpSerG 435

735 rGtYLeuAspLysPhePheAlaLeuValAsnAlaLeuAspGluHisLeu 751
|||||
2177 CCTACCTGGACAAATTTCTTACACCTGGTGCAGCTGCTGGAGAGGAGACC 2226
752 PheProValArgIleGlyAspMetArgIleMetGluAsnLeuGluAs 768
|||||
2227 TTCCCATTCCTGGCTCAGGACACTGCTGAGCGGAGGCAACGTGAGCA 2276
768 nGluLeuLysSerSerIleSerAlaLeuAsnSerSerGlnLeuGluProV 785
|||||
2277 GGAAGCTGGGGGCAAGTCTTGCAGCACTGGCTGGCCAGCCCGAGACCC 2326
785 aValArgPheLeuHisLeuLeuAspLysLeuIleLeuValIle 801
|||||
2327 TTGTGGCCTTCTCCACACAGCTGCTGAGCAAGCTGCTGCTGCTGCTC 2376
802 ArgProProValIleAlaGlyGlnIleValAsnLeuGlyGlnAlaSerPh 818
|||||
2377 AGGCCCCGATCATCATGCGCAGATGTGAACCTGGCCGCTGAGCCTT 2426
818 eGluAlaMetAlaSerIleIleAsnArgLeuHisLysAsnLeuGluGly 835
|||||
2427 TGAAGCAATGGCCATGTAGTACGCTGTTCACCGGAGCCTGGAGGAGAG 2476
835 snHisAspGlnHisGlyArgAsnSerLeuLeuAlaSerTyrIleHisTyr 851
|||||
2477 CCCAGAGATGCCCGGCTCAGCTGCCACAGCTGGCTACGTCACATAC 2526
852 ValPheArgLeuProAsnThrTyrProAsnSerSerSerProGlyProG 868
|||||
2527 GCTTTTCGCTCTTGGGCACTGAGCCAGCCCTCCGATGGGGCCCTT 2574
868 yGlyLeuGlyGlySerValHisTyrAlaThrMetAlaArgSerAlaVal 885
|||||
2575CCAGTACATGTCAGGCTGCCACACTGGCCGCTGCTGCTC 2617
885 rGProAlaSerLeuAsnLeuAsnArgSerArgSerLeuSerAsnSerAsn 901
|||||
2618 GCCCGCAAGCCTTACCTGCGGCGCTTCCAAAGACATCAGCAGCAGCAAC 2667
902 ProAspIleSerGlyThrProThrSerProAspAspGluValArgSerI 918
|||||
2668 CCTGACTTCGCGCTGGCCCTGCTGATGACGAGGTTTCCCGCAT 2717
918 eIleGlySerLysGlyLeuAspArgSerAsnSerTyrValAsnThrGlyG 935
|||||
2718 CCTGGCCAGC..... 2727
935 LyrProLysAlaAlaProThrGlySerAsnProSerProSerAlaGluSer 951
2727 2727
952 ThrGlnAlaMetAspArgSerCysAsnArgMetSerSerHisThrGluTh 968
2727 2727
968 rSerSerPheLeuGlnThrLeuThrGlyArgLeuProThrLysLysLeuP 985
|||||
2728AAGCTGC 2734
985 heHisGluGluLeuAlaLeuGlnThrValAlaCysSerGlySerValArg 1001
|||||
2735 TTCACGAGAGAGCTGGCTGCGAGTGGTGTAGCAGCAGATGCCCTAGCC 2784
1002 GluSerAlaLeuGlnGlnAlaTrpPhePheGluLeuMetValLysSe 1018
|||||
2785 GAGGCCATCTCCACAGACCCGTGTTCTTCCACATGATGTGAAGAG 2834
1018 rMetValHisHisLeuTyrPheAsnAspLysLeuGluAlaProArgLys 1035
|||||
2835 TTATGGCGCTGCACCTGCTGCTTGGCCAGGAGCTAGACACACCCCGCAAGC 2884

1035 eArgPheProGluArgPheMetAspPheIleAlaAlaLeuValSerThr 1051
|||||
2885 TGGCTTCCCGGAGCGCTTCTGTGACGACATCATGCTGTTGGGCTCT 2934
1052 IleAlaSerAspIleValSerArgPheGlnLysAspThrGluMetValG 1068
|||||
2935 GTGGCGCTGGAGGTCAACCCGTGTCCACAGAGATGTGAGTGGCCGA 2984
1068 uArgLeuAsnThrSerLeuAlaPhePheLeuAsnAspLeuSerValM 1085
|||||
2985 GCACCTAACGCCACGCTGGCTTTCTTCTCTCAGTACATCTTCTCTG 3034
1085 eAspArgGlyPheValPheSerLeuIleLysSerCysTyrLysGlnVal 1101
|||||
3035 TGAGCCGGGGCTTTGCTTTCAGCCTGGTCCGGGCCACATACAGAGG 3084
1102 SerSerLysLeuTyrSerLeuProAsnProSerValLeuValSerLeu 1118
|||||
3085 GCCACGGGCGCTCAGTCTGCTCCATATCCAGACGCCCTGCTGACCTG 3134
1118 gLeuAspPheLeuArgIleIleCysSerHisGlnHisTyrValThrLeu 1135
|||||
3135 CATGGAATTCACCCGCAATCTGTGACAGCCAGACATACGTACCTCA 3184
1135 snLeuProCysSerLeuLeuThrProProAlaSerProSerProSerVal 1151
|||||
3185 ACCTCCCTCTGTCGCTGCTGACCTGACCTGACCTGCTGCTGCTG 3234
1152 SerSerAlaThrSerGlnSerSerGlyPheSerThrAsnValGlnAspG 1168
|||||
3235 TCTCCACACCACTCCAGAGCTCCACCTCTCCAGCCACACCCCGGAGCC 3284
1168 nLysIleAlaAsnMetPheGluLeuSerValProPheArgGlnGlnHis 1185
|||||
3285 CAAAGTACCAAGCATGTTCGAACTGAGTGCACATTCGCCGACACACT 3334
3335 TCTAGCTGGGCTCTGCTGACGAGCTGGCAGCTGGCCCTCGAAGCTGAG 3384
1202 AlaGluGlyLeuPheGlyLeuHisLysValIleAsnMetValHisAs 1218
|||||
3385 GCTGAAGGGCATTCCTGTTGCACAAAGACGATCATGCTGTGCACAG 3434
1218 nLeuLeuSerSerHisAspSerProArgTyrSerAspProGlnIle 1235
|||||
3435 CTGCTATGTGGCCATGACACTGACCCCGCTACGCGGAGCCACTGTGA 3484
1235 ysaAlaArgValAlaMetLeuTyrLeuProLeuIleGlyIleIleMetGlu 1251
|||||
3485 AGGCTGTGGGCGGAGCTGTACCTGCACTGCTTTCGATGGACGGAT 3534
1252 ThrValProGlnLeuTyrAspPheThrGluThrHisAsnGlnArgLys 1268
|||||
3535 ACCTTCCACAGGCGTCAATGCTGTGAGAGGCCAGTGCAGCGTCAAG 3584
1268 gProIleCysIleAlaThrAspAspTyrGluSerGluSer.....Glys 1283
|||||
3585 ACTGGCCTCAATGCTGTGACACACAGAGAGGAGGAGCATTTGCGG 3634
1283 eMetIleSerGlnThrValAlaMetAlaIleAlaGlyThrSerValPro 1299
|||||
3635 GTACATCAACACCTCTGTGGCATGGCCATTTGCTGGTGGCCGCCA 3681
1300 GlnLeuThrArgProGlySerPheLeuLeuThrSer.....Th 1312
|||||
3682GCCCTGCTCTCCGGGCCAGCATCTCCAGGGGCGACACAC 3722
1312 rSerGlyArgGlnHisThrThrPheSerIleArgLysSerArgSerLeu 1329
|||||
3723 GGTCTTCGGCAGAGCTGTGCTCTGTGATCAAGCGGACCTTGC 3772
1329 euIleCysLeuLeuThrValLeuLysAsnAlaAspGluThrValLeuGln 1345

4652	TCGCAAGACGAGGTCCAGGACCTGATGTTTCAACCTGCACATGATCTCTGACG	4701
1646	AspThrValLysMetLysGlnHisGlnGluAspProGluMetLeuIleAsn	1662
4702	GACACCGTGAAGATGAAAGACACCCGAGGAGACCCCTGAGATGCTCATGCA	4751
1662	pLeuMetTyrArgLleAlaLysGlyTyrTrpLinhSerProGlu...ArgL	1678
4752	CCGCATGTACAGATATGCCCGGGGCTACCAAGGAGCTTACCGGACCTTCGGC	4801
1678	euthTrpLeuGlnAsnMetAlaGlyLysHisSerGlnLysSerAsnHis	1694
4802	TGACCTGTGTGAGACATGCGCGGGAACGACGCGGAGCTGGGCAACGAC	4851
1695	AlaGluAlaAlaGlnCysLeuValHisSerAla	1705
4852	GCGAGGCGCGCCACATGCTGTCACGCGGCGC	4884
seq_name: gb_ov:AY016023		
seq_documentation_block:		
LOCUS	AY016023	167215 bp DNA linear VRT 01-FEB-2002
DEFINITION	Spherooides nephelus alpha globin gene cluster, complete sequence.	
ACCESSION	AY016023	
VERSION	AY016023.1	GI:18463963
KEYWORDS	.	
SOURCE	Spherooides nephelus.	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Euxariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Acanthomorphi; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Spherooides.	
REFERENCE	1 (bases 1 to 167215)	
AUTHORS	Flint,J., Tufarelli,C., Peden,J., Clark,K., Daniels,R.J., Hardison,R., Miller,W., Phillips,S., Tan-Un,K.C., McMorrow,T., Frampson,J., Alter,B.P., Frischauf,A.M. and Higgs,D.R.	
TITLE	Comparative genome analysis delimits a chromosomal domain and identifies key regulatory elements in the alpha globin cluster	
JOURNAL	Hum. Mol. Genet. 10 (4), 371-382 (2001)	
MEDLINE	21096913	
PUBMED	1157800	
REFERENCE	2 (bases 1 to 167215)	
AUTHORS	Flint,J., Tufarelli,C., Peden,J., Clark,K., Daniels,R.J., Hardison,R., Miller,W., Phillips,S., Tan-Un,K.C., McMorrow,T., Frampson,J., Alter,B.P., Frischauf,A.M. and Higgs,D.R.	
TITLE	Direct Submission	
JOURNAL	Submitted (06-DEC-2000) MRC Molecular Haematology Unit, Institute of Molecular Medicine, John Radcliffe Hospital, Oxford, Oxon OX3 9DS, UK	
FEATURES		
source	Location/Qualifiers	
	1..167215	
	/organism="Spherooides nephelus"	
repeat_region	/db_xref="taxon:39110"	
	1510..1530	
	/note="simple repeat"	
	/rpt_type=tandem	
	/rpt_unit=c	
repeat_region	2315..2350	
	/note="C-rich, low complexity"	
repeat_region	2643..2663	
	/note="simple repeat"	
	/rpt_type=tandem	
repeat_region	3620..3676	
	/rpt_unit=ca	
	/note="simple repeat"	
	/rpt_type=tandem	
repeat_region	4231..4284	
	/rpt_unit=ta	
	/note="simple repeat"	
repeat_region	4638..4678	
	/rpt_type=tandem	
	/rpt_unit=ca	
	/note="simple repeat"	
repeat_region	4638..4678	
	/note="simple repeat"	

```
/rpt_type=tandem
/rpt_unit=tg
repeat_region
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=ca
repeat_region
/note="AT-rich; low complexity"
8334..8354
/note="AT-rich; low complexity"
9950..9970
/note="AT-rich; low complexity"
10777..10802
/note="AT-rich; low complexity"
join(<10872..11080,11797..>11973)
/gene="unkempt"
/product="unkempt"
<10872..>11973
/gene="unkempt"
join(<10872..11080,11797..>11973)
/product="unkempt"
CDS
/gene="unkempt"
join(<10872..11080,11797..>11973)
/gene="unkempt"
/codon_start=1
/product="unkempt"
/protein_id="AAL73045.1"
/db_xref="GI:18463964"
/translation="YIKERFTEOCPLFLQHKCTQHREPTCFNNHFLNQRRRPIRRRD
GTFNYSDDVCTKXDETGTGLCPGDGCPYLHRTTGDERKYLHRYKKTGTGCIHETDAR
GHCVKNGIHCAFAHPDRLRPVYDI"
17726..17763
/note="A-rich; low complexity"
19346..19371
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=c
20068..20121
/note="AT-rich; low complexity"
complement(<21113..>22747)
/gene="IGFALS"
/product="IGFALS"
complement(<21113..>22747)
/gene="IGFALS"
complement(<21113..>22747)
/gene="IGFALS"
/codon_start=1
/product="IGFALS"
/protein_id="AAL73046.1"
/db_xref="GI:18463965"
/translation="ACARGCTCOHDYSLNLYCSAQNITQVPPDMPSTHSLMIDY
NLASLPAASKFDVLEFNLDSGQVLTDPQALRGLSLAHILERNIRVLPGLI
FQNTPKLASLSLHNNQSLRIDRLPAQLSHMWLNLGRNSIAVLPETFODLQGLKEL
ILVGNRLAYLQPOLFQNLVELKELIDLSGNHLKYIKANVPVKLTKLOKLYDAQOIMTV
VPRAFVGMKSLRMVLELNNRLSLHEDTFMGLYLLHVRVSNNSIGIKRPTREDSQ
LEELRLSHNRIRVIGEKVPEFELRLELEHNNIOEAKVGSFTGSHAVINLSSC
PHILPDVFRGLSKLSLHIDRGLITRVTTQASGSLSLRLPLQHNINISVEROSYV
DLVGLIDLSFNKLELVGTQTSGLNLTLLISNDCKRHLQNGSRLLPLRUYD
LRANELTSIVPEEFAEGLEKLLSGNRWRCOGALPLRNYSLRALVYPRQVETHABE
EPDITITVYNNITCSPPRLTGODLRVEGVSFVQSC"
24332..24356
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=gga
24525..24554
/note="GC-rich; low complexity"
24862..24895
/note="AT-rich; low complexity"
26315..26340
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=ca
26915..26951
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=cgg
```

```
repeat_region
27240..27272
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=tg
repeat_region
27487..27507
/note="AT-rich; low complexity"
28056..28080
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=tg
28159..28189
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=ccg
30084..30122
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=cag
30217..30260
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=tg
31181..31214
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=gq
33149..33186
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=tg
33298..33379
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=ca
33966..34015
/note="T-rich; low complexity"
35457..35500
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=ca
35965..35993
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=ctc
37605..37637
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=tg
37782..37804
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=ca
40208..40383
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=ctca
40386..40562
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=tcta
41620..41645
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=ca
41968..41994
/note="simple repeat"
/rpt_type=tandem
/rpt_unit=tgg
44729..44773
```

/note="simple repeat"
/rpt_type=tandem

alignment_scores:

Quality: 4986.50 Length: 5470
Ratio: 2.677 Gaps: 90
Percent Similarity: 34.059 Percent Identity: 26.362

alignment_block:

US-09-737-246-2 x AY016023/rev ..

Align seg 1/1 to reverse of: AY016023 from: 1 to: 167215

```
49 ValProLeuThrcgluAlaValAspProValAspLeuGluAspTyrLeu1 65
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
134857 TTGCCCCCTACACAGAGGTGGAGCGGTGATTTTGAGAGATTATGTGAG 134808

65 ethrHISProLeuAlaValAspSerGlyProLeuArgAspLeu1eglup 82
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
134807 CAGCCATGCTCCTGGGGGTGGAGCCTGTCCCTCAGAGAGCTTGAGGT 134758

82 heProProAspAspIleGluValTyrSerProArgAspCysArgThr 98
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
134757 TCCTCCCAAGACGAGCTGGAGCTCTCCACCTGGACAAAGAGTCACAAAC 134708

99 LeuValSerAlaValProGluGluSerGlu..... 108
|||:::|||||:::|||||:::|||||:::|||||:::|||||
134707 CTGGAGCCCCCTGCTGCTGATGAGAGAGTAAAGTTTGTCCCTGTGCC 134658

108 ..... 108
134657 AACACAATGCTCACACACACTGTCTTTCGCTTTTAAACAACACACA 134608

108 ..... 108
134607 CACACACACACATACACACAGATGATGATTACATATTGGAGTGAGAG 134558

108 ..... 108
134557 TGAACATACAGAACAAAGCCTTTTGAAGTGGCTGAAATAAACAAAC 134508

108 ..... 108
134507 ACAGTTACTGTATTAAGCAATTTCTGCTCAATTATTCATGTTGTGTC 134458

109 .....MetAspProHisValArgAspCysIleArgSerTyrThr 121
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
134457 TGTGTAGCTCACTGATCCCAAGAGTGGCGACACCTTGTCACTATACACA 134408

122 GluAspTyrAlaIleValIleArgLys..... 130
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
134407 GATGACTGGCTCATCTTCAAGAAAGTAAAGTACATGTGAGGACGACA 134358

130 ..... 130
134357 CCGTTAAATGTGTTCCTCTGCAAGAACGCCCTTTTATCTGTTGATTG 134308

131 .....TyrHisLysLeuGly 135
|||:::|||||:::|||||:::|||||:::|||||:::|||||
134307 CTTCGTTATTAACCTGGCTCTCTTTTATGTAGTATACACGATTAACAGC 134258

136 ThrGlyPheAsnProAsnThrLeuAspLysGluLysGluArgLysGly 152
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
134257 ACCACTTTACAGCCTCACAGCTCAGAGACAGAGAGAGGACAGCGAGG 134208

152 yLeuProLysGluValPheGluSerAspGluAlaProAspGlyAsnSerT 169
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
134207 GCTGGTCAAAAGACACTTGTAGCTGGATGTGCG.....GCCACTGAGC 134164

169 yTGlnAspAsp..... 172
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
134163 ACCAGGACGACGAGTAGCTAATGTATCTGATCGAGTGTGATCCGAGT 134114
```

```
172 ..... 172
134113 CCATACAGTGAATTTATTTGACCACACCGTATTATATATTCATTAATG 134064

172 ..... 172
134063 TGGCAGGAGAAAGCATTCAGAGTTTTTTTTTCTTTTATATATGAT 134014

172 ..... 172
134013 AGCTGACAAAGACACCAACCACTGAAAAAGCTGTGTCTCTGATGG 133964

172 ..... 172
133963 GAAATAGAAAAGACAGACTACTAGAAATGATATACCTGCTCATTA 133914

172 ..... 172
133913 TCTCCACATTCCCATTTATATCATATTAAATAGTAATAGTTAATGGGT 133864

172 ..... 172
133863 CTGTATCATATTACTCTATAAAACATCTCTGCAGACATTAAAGGGGC 133814

173 .....GlnAspAspLeuLysArgArgSerMetSerIleA 184
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
133763 CCTCTTACGATCCACAGAGTACGCAAAAGCGGCTGACGTACGCTCG 133714

184 sPAspThrProArgGlySerTyrAlaCysSerIlePheAspLeuLysAsn 200
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
133713 AAGAACTCTCTGAGGAGCGTCCACGACATTTGACCTGAAGAAC 133664

201 SerLeuProAspAlaLeuLeuProAsnLeuLeuAspArgThrProAsnG 217
|||:::|||||:::|||||:::|||||:::|||||:::|||||
133663 TCCACCCCAAGTCTCTCTGCGACGTGTCTTAAACGACACGGCTCGGA 133614

217 uGluIleAspArgLysAsnAspAspGlnArgLysSerAsnArgHisLysG 234
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
133613 GGACATGAGACACCGCAATACCGAGGCGGCCACAGAGGCGCCACAGCTG 133564

234 IuLeuPheAlaLeuHisProSerProAspGlu..... 244
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
133563 ACCTGTGGGCTTTTACCACCCCTGATGAGGTTTCTCAAAAGCAATTT 133514

244 ..... 244
133513 GTGATCTTGAATATATGATTTTATTAATACACTGACATGACCAAT 133464

244 ..... 244
133463 TCTAACATATTTCTCAGCAACACACATTTAATCAAAAATTAATTAAT 133414

244 ..... 244
133413 AATTGATGATTTCACACTGGAATATGCGGCTCTGACAGATGAATGTGT 133364

244 ..... 244
133363 GATTTAAGCCTGTGTGTTATTTTGAATAGCTTACATTGACCTATGAACA 133314

244 ..... 244
133313 TTTATGCACTGTTTAAATAATGATGACACTTGAATGTTTGTGATGT 133264

244 ..... 244
133263 AAGTGATTTTGTCAAGCTCTGGGTAAAGTAATATGCTTTGTCTCAAG 133214

244 ..... 244
```


477 VASPARgLeuSerAspGluAspLeuTyrLysPheLeuIlaaspMetArg 494
|||||
131381 GGAACGACCTAACTGAAGAGACCTTTATTAAGTTCTGGCAGATATGCCA 131332
494 rProSerSerValLeuArgArgLeuArgProIleThr..... 506
|||||
131331 GACCGCTTCTGCTGGCAGCAGTGAAGGCGCTGCACAGATAGCATTTAA 131282
506 506
131281 GACTTACTTGTACAAAGATAATGTCTAACATGTGATGTGTTGTC 131232
506 506
131231 CTCTCTAACTTAAGATGCACGTTAAAGTGTTAACTCCATTTTCT 131182
507AlaGlnLeuLysIleasp 512
|||||
131181 TCTTTTGCTGTGTGATTTATTTGACCGCGACGCCAGTTGAGATTTGAT 131132
513 IIEserProIaProGluAsnProHisTyrCysLeuThrProGluLeu 529
|||||
131131 ATCTCTCCAGCTCCGAGCTCCGCTCATTACTGTGTACACAGAACTGCT 131082
529 uGlnValLysLeuTyrProAspSerArgValArgProThrArgGluIleL 546
|:::|
131081 TCATGTAAACCCCTACCCAGACCCAGCGGTGCGCCCAACCAAGAAAGTC 131032
546 euGluPheProIaArgAspValTyrValProAsnThrThrTyr..... 560
|||||
131031 TGGAGTTCCTCTGCTATGTGTACAGCCACATACCTCCATATGATGA 130982
560 560
130981 GAGCCCCCAACTCAGAAATCATTTATATCTTCTGTCAATATCA 130932
560 560
130931 TTAATTTTATGCTGCTTATTTTCTGTATTAATTTCTGATATATTAGAG 130882
560 560
130881 GATAAACACAGATAACCTATGTGAAAAACATATTTATATACATTCT 130832
560 560
130831 AATAGAAATAACATTCTCAGTAATTTGCATGTTAGTCAAGTTTAATA 130782
560 560
130781 AAGTAACAATTGTTTCAAGTTAAGCTAAGTTCACAAAGTACAGCACAA 130732
560 560
130731 ACTCCATCATCAGGTCAACAGATAAAGAACCTGTAGTAAGSTCAT 130682
560 560
130681 GTCAAGTTCAGCTTACATACATGTGAGGAGAGAAAGAGTGCAGTCAATG 130632
560 560
130631 TTGCAAGTAAAAAAACCTGTGGCAATCTTAAATGGCCATGCTGTGC 130582
560 560
130581 CTTGCTCTTGCAGATGCAAGAGAAACAGATACAGATTAGAGGAT 130532
560 560
130531 ATGAATAAACACGGGCGGTTTAAACAAATATCTTACGTTTCTGTAG 130482

560 560
130481 TTTAGTAGCTTTTCACTGTGACTTTGATGCTCTCTTTTGAACGTAA 130432
560 560
130431 TCAGTCTGTAGTCTCTCAAAACTTCAGTGTGTCCTCGCTTAAGTC 130382
560 560
130381 TAGACGGGAGAGAACAGTTTCCATATCATTTTGTGTGTATGCTG 130332
560 560
130331 TATTGCTCTAGTATTTATTTACAGAGAAATTCGTTACTGCAGCTCTTT 130282
560 560
130281 GATGCTGCTGTGGGTACATATTTATGAGATGTGTCTCCAGCCTCTTTA 130232
560 560
130231 CATTTGTCTGAGCTCCACTTCTTCTGCTGCTCAGGTTGGCCAGCTGTCT 130182
560 560
130181 TTGCTCACCTCGGCTGTGAGCTGTGCTGGCCATCTGGGTCTGTTGAAA 130132
560 560
130131 GAGCATCTGTCCATGTATTTCTCTCTACAGCAGCAATTCACGCGAGCTG 130082
560 560
130081 GCACCGCGCTGACCCCGTATTAACATGTGGCTTTCTAATGACACACTCTC 130032
560 560
130031 AGATGAAGGCTCTGCTATATCTGCAAACATCATGAAAAACCAATGTACAG 129982
560 560
129981 CCAAAATCACCAAAAGTCTAGAAAAACATTTATCCCAAGAGAACAG 129932
560 560
129931 TGCTTTATTCAGATTTTGTAGATATGACCTGTGTTTCTGGCCCTCTCG 129882
560 560
129881 CATGTTTCGGTCTGCAGTTTGATACAGTGTGTGTCTCTCTCTCTCT 129832
560 560
129831 AGCAGCAGACCCCTTGAGTAGCCAGACACATGTTGTGACATATTTCT 129782
560 560
129781 GCCATTCATTACCTGTAGACAATTTATACGTGTTAGCCGTGTTTTCAT 129732
560 560
129731 TCTCTGAGCTTCTCTAGCATGTCTCCCAAGCTTTGAAGTATCTTTCA 129682
560 560
129681 TGGTTCCTCAAGCTGTGCTTATTTCTGTTATCTTGAGCTCGGTGCTTC 129632
560 560
129631 AGAAATGTATTTAGAGGAGTGCCGAACGTGTATGAATCCAAACGTGAGCA 129582
560 560

129581 CGTTGACCTGCTGCTGGTGAACATGTTTTCCTGCTGCTTTCCTC 129532
560
129531 ACTGGACCTGCACGGGTTCTCTTAACCCCGACCAACCGATCAGACA 129482
560
129481 GAGACCTCAGATCAGATTCATAGTATTAATCCCTCAGTGTAGTCCAGCTG 129432
560
129381 TAAAGTATTCGTAATAAGTCCCATCTGTTTAAGATCCAGTGTGTGCT 129332
560
129331 GGGTCTGTGTAATTTGGAGGGTAGTCAGTGCATCTTATATTGGATGNG 129282
560
129281 TATATGATAGTAATAATTAAGTCTAGCTGTGATGTCATCAGTGACCTT 129232
560
129231 ACTCTCTTGACCTGCCCTAATGTTTAACCCCTGAGATGATGATCTTGA 129182
560
129181 CATTTAGTCAACAACACATGTGACACATTACTGTATTTTGTCTTGA 129132
560
129131 CACAGATCTTGGTGTGTGTTTTCATGCCACACTGCTGTGTAATGTTGT 129082
560
129081 AAAAGATTTCTAGTGTGCCAACAATCTTAACTCATATCAAAAGAAATTA 129032
560
129031 CATCAGCTGTGCCAAAGTTACTCAGATGTTGACTTTAGTTGTGGAAT 128982
560
128981 TACGGTACTTTGACTCCACAACATCCAAATGCTGGGTCACTCCTGTTA 128932
560
128931 CATAGATATTGCACATTCATGAATACTTCTTACTCTTCTAAGATT 128882
560
128881 GTTACTGATAGACAGTGGTAGATGCTGTTGACAACAGAATGTGCCAAT 128832
560
128831 TTGCATTCATATTTAAGATTTGGACAAGATCTTCCTTTCAAAATCA 128782
561 Arg..AsnLeuLeuTy
128781 GCATTTCAAAAAATTTATATAAACCATGCTTTCGACGAGACTTCTCTTA 128732
565 rIleTyProGlnSerLeuAsnPhelaAsnArgInglYSerAlaArgA 582
128731 TGTATCCCAAAAGCTTCAACTCAGTAGTCGTCAGGGCTCAGTAGAGA 128682
582 snIleThrValYsValGlnPheMetTyrgLYsAspProSerAsnAla 598
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

128681 ACATCGCTGTCAAGTCCAGTACATGGCAGAGAGACCCCACTCAAGCT 128632
599 Met.....
128631 TTGCAAGGTGACGACAAAGTCAAGGACTACCAAGTGAATAAATAACATGTT 128582
599
128581 CTAGAAAAACAGCACCGAATCTGTATTTGAGAAATTTACCTGGCTTTAG 128532
599
128531 TAAATTCGAGTAAGTAAAAAACAATCCCTAATGATTTTGTATT 128482
600 Pro..ValIlePheGlyYsSerSerCysSerGluPheSer 612
128481 TTTATTTTCACGTAGTCACTTTTGGGAAGTCAAGTTGTCTGAGTTCCACA 128432
613 LysGluAlaTyThrAlaValValTyHisAsn..... 623
128431 AATGAGCATATCAGTCCCATCATCTACACACAAAGTAAGCTTTCAGA 128382
623
128381 CTTAACCTGCTGTTAACCTCCTTGTGTTTCTTAATTCGTGTTAA 128332
624 Arg..SerProAspPheHisGluGlu 632
128331 CTCTCTTTTGTGTTAAAAAAGAGTCCCTGAGTGTATGAGAGACA 128282
632 IeLYsValLYsLeuProAlaThrLeuThrAspHisHisLeuLeuPhe 648
128281 TGAAGATGAAGATCTCGCAATCTGACTGACAAATCACCACTCTGTTTC 128232
649 ThrPheTyHisValSerCysGlnGlnLYsGlnAsnThrProLeuGluTh 665
128231 ACCTTTCACACATCAGCTGCCAGCCCAACAAACCTCTCTAGAGAC 128182
665 rProValGlyTy..... 669
128181 TCCTGTGGGTACACGTGAGTTCACCTAGCTACATTTTATACT 128132
669
128131 CCCACCCACAGCCCGACGTTGATATTTCTGTTGTACATGCCATCTT 128082
670 Thr..TrrIleProMetLeuGln 676
128081 AAAAGATCTGCTGCTTTCTTTCGACTTGTAGTGATCCCTGATGTCAG 128032
677 AsnGlyArGLeuLYsThrGlyGlnPheCysLeuProValSerLeuGluTy 693
128031 CACGGCCGACTGGAACAGGCTCCTTCAAGTTTACCTGTCTCAAGTGAGAA 127982
693 sProGlnAlaTySerValLeuSerProGluVal..... 705
127981 GCCCCACCGAGTACTCTGTCTCACCCCGATGTAGAACAGACAGCTGGG 127932
705
127931 GATATCAGTTTGCCGAAAAATGACCTTGGCGTTTGTACTGGCTGTTAA 127882
706 Pro..LeuProGlyMetIleTyTrpV 713
127881 TGACATGATTTGCTGCTGTGATGATGTTTCAGCTTCCAGGATGAAGTGGG 127832
713 aLAspAsnHisLYsGlyValPheAsnValGluValAlaValAsnSer 729
127831 TGGATATTCACAAAGAGGTGTTAGTGTGAAGAGACAGAGCTCTCTTCG 127782
730 ILeHisThr..Gln..... 733
127781 GTTCATACTCAGGTACAGACGCTAAATAATTTGCTTTCATTTCTATCAT 127732

```

733 ..... 733
127731 CTGACGACGATCATTTCAAGACATCCATTTGATCTTAATGTTATAGTAAA 127682
734 ..... AspProTyrLeuAspLysPhePheA 742
127681 TATTATAGTAGTCTGTAGCTAGTACCCCTTCTCGATAGTACTTCTTCA 127632
742 lAlaValAlaSnAlaLeuAspGluHisLeuPheProValAlaArgIleGlyasp 758
127631 CCCTGCTGATGTCTCGAAGAGTACTCTCCCTCCGCTCGAAGAGAT 127582
759 MetArgIleMetGluAsnAsnLeuGluAsnGluLeuLysSerSerIleSe 775
127581 GTGATCATTTACTGAGTCACAAACATGGAGGCGAGGTTAAAGACCATGTGT 127532
775 rAlaLeuAsnSerSerGlnLeuGluProValValArgPheLeuHisLeuL 792
127531 GGCACGTGAGGGGCCCTGTGTGATACCTGTGTGACGCTTCTCCACAGC 127482
792 euleuAspLysLeuIleLeuLeuValIleArgProProValIleAlaGly 808
127481 TGCTCAACAGATCATTTCAACTCATCGTGTTCACACCGTCATCGAGGC 127432
809 GlnIleVal..... 811
127431 CAGATTGGTAAATGAGCTCTGAAGCACCGTCGCTGTCACTGCTCAACAA 127382
811 ..... 811
127381 TTGTGATTTGGAGTGTGCTCTCTGTGAACAAAGGTGAAGGGCTCATTTGA 127332
811 ..... 811
127331 ATAGTGGGTTAATAGTTGATGTGATTAATGATCAAAATTACTGATGAA 127282
811 ..... 811
127281 ATAATGACATTTTGTCTCATACACATTAATAAACACTGACTCATGTGAT 127232
811 ..... 811
127231 TGTAAACCATGTGTTTAAATATCCTACATCTTGATTTCCACATC 127182
812 ..... AsnLeuGlyGlnAlaSerPheGluAlaMetAlaSerIleI 825
127181 CTGACAGTCACCTTGGCAGGAGCTGCTTGAACATATGGCTTTTACTGCT 127132
825 eAsnArgLeuHisLysAsnLeuGluGlyAsnHisAspGlnHisGlyArgA 842
127131 GAACCGAGACACAAAAACCTGGAGGAAACCGAGCACCATCATGTGTCGA 127082
842 snSerLeuAlaSerTyrIleHisTyrValPheArgLeuProAsnThr 858
127081 ACAACCTGCTGGCTTC.TACATCCAAATACTGCTTGGCTGGCCACTGCT 127033
859 TyrProAsnSerSerSerProGlyPro..... 867
127032 GAACCATC...AGTGGCCCCAACAGAGTGAGACCTGCGACACATGTGCA 126987
868 .....GlyGlyLeuGlyIleSer..... 873
126986 CAGCAGATTAATTGGTTGGATTGCTGGCAGTGTCTGAGACCGGGCCGCT 126937
873 ..... 873
126936 TGATCTGCCTTATCTTCTGTGGCTCATACAGTTTGAGAGAAGACC 126887
874 .....ValHisTyrAlaThrMetAlaArgSe 882
126886 AGTTCACTGCCATCATGATGCCGTGCGAGTAGCGAACCATATATCGAGGC 126837

```

```

882 rAlaValArgProAlaSerLeuAsnLeuAsnArgSerArgSerLeuSerA 899
126836 AACTGGACGCCCAAGACGACCTGTGACACCTGTCCCGCTCCAAAGCATGACGA 126787
899 snSerAsnProAspIleSerGlyThrProThrSerProAspAspGluVal 915
126786 ACTCAACCCAGACACCTGGCCAGTTCACACAGTTTCTCCGATGAAGAGTG 126737
916 ArgSerIleIleGlySerLys..GlyLeu..... 924
126736 CAAGGATCATAGGAGCAAGGTGAGCTCTTTGTGCTAATGAAGCTTTT 126687
924 ..... 924
126686 GTCTTTTATTACAGCTTAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 126637
924 ..... 924
126636 GTGTGTGTGTGTTAAGATTGAGAACATTAATTCACAGTAGAATCTGTTA 126587
924 ..... 924
126586 CATACTGTGTAAGAGAGTGAAGATTTCTCTCGTTTTCACAAACAAACACAG 126537
924 ..... 924
126536 TTTTACACAAAGTAAGTTTTTTGGGGGTATTCTTGCTGCCCAAAA 126487
924 ..... 924
126486 ACTGAATGTGAACCTTCTGGGCTAGGCGCTTCCACACAGTGTCTTAC 126437
924 ..... 924
126436 ATGTGCGCTGCGCCTTGTGTCAAGCTGTGTCTCTGCGCAATATTA 126387
924 ..... 924
126386 TCCACCAATTATTGCCCATTCGCTCTTGTGATTAATATAATTCCTTAAT 126337
924 ..... 924
126336 TTGGGTTAAGTGTGAGAGACGAGCTGAGACACTCCCATATCAGGGAATG 126287
924 ..... 924
126286 ACATCATTTAGTTGACCGCCCGGTGAGCTGATTCATTTCTCGCTTCT 126237
924 ..... 924
126236 TTTAGACTATGTGTTAATATGTAATTTTTTTATGAAGTGCAATTTATGCTT 126187
924 ..... 924
126186 TGTGTTTGAACATGTCTTGAGCCCTTGATTTCTGTGCCAGTGTGAG 126137
924 ..... 924
126136 ATTCTGTGTGTTTAAGCTTTATTTCTGTGCGCCCGCCCGCTTTTCT 126087
924 ..... 924
126086 TTGTTTTTCACACATTTATCTCCAAATAATTAATTACTTGTGTGTATTAAG 126037
924 ..... 924
126036 CTTCCTTCACTACTGTATACCTTGTGCTGCGCACTGTGTGTTGTGT 125987
925 .....AspArgSerIleAsnSerTyrVal 931
125986 TTGTCTTCGGTGCTTATGCGACAGGCAATTAACCGCTGCCACTCTCTGGGTT 125937
932 AsnThr.....GlyGlyPro..... 936

```

125936	AACCTGCTTATGGCCCCCGGGGTCCAGATCATTCGTTGCGCGAACC	125887
937LysAlaIaIaProTrpGlySerAsnPro....	945
125886	CAACTCAGCTGTGATCTGTGAAGCAGGTGCCAAATCCCTCTCTCGCTCCT	125837
945	945
125836	CCTAACCTACGCGTGCCGCACCGTCTMAACAACACTCTTGAGTCATTGT	125787
945	945
125786	TGTGTGCAATAAGCTTCCATCTGCTTTTCATGCTGCCACTCCTTCTTAATC	125737
945	945
125736	TGTGAATACACTTATCTGTGTTCCACCATTCATCCCTCACACCCAATAA	125687
945	945
125686	CAGCCCACTGTGTGTGAGTGCCTGCTGTCAGAGTTTTTACTAGC	125637
945	945
125636	CTCACACTATCGTACAGAGGTCTGTGTCACCCACTAAATTTCCATTCAAT	125587
946SerPro.SerAlaIaIaSerThr	952
125586	CCATTTTCACTACATTTTGTCTGTGTGTTCACCTCACACTCACACCATGT	125537
953	GlnAlaMetAspArgSerCysAsnArgMetSerSerHisThrGluThrSe	969
125536	CAGACAAGCAGCAGCAGCTGCATTCGATCTGTCTTCCTGGACAGCAT	125487
969	rSerPheLeuGlnThrIleuThr.GlyArgLeuProThrLys.Lys....	983
125486	GGCCCTTGTTCAGTTCCACAGAAGCATATGCCAAGAAGGTAAACGTGT	125437
983	983
125436	CTTCTACTATACAGAATACAGAGTGTCCAACTCAAAGGATGCATTTT	125387
983	983
125386	GTGCTCTCATGATTCATTCATTTTCACTGTCATCATCATGCTACATCA	125337
983	983
125336	CCATCATATGAGTTAAAATGCTGACACGTGATCTAAACATTTTAAACCA	125287
983	983
125286	AGCATATATTATTATTTCAATTATTCATTTTAAAAATATAATACTTAAT	125237
983	983
125236	AACACTTAATCTCACAAACAGCTGATTTGTTTAGATGTTTCTGTTTTC	125187
983	983
125186	AATTTAAATCTAATTCATAATTTTAAACATTAATATATAGACAAATTC	125137
983	983
125136	TAAAGTAAAGAGATGTGTAAAACTTTCAGTGTGTTTGTTCCTGTA	125087
983	983
125086	AATGTGATTACAGTTTATATGATGGAATATTTAGATTTTGTAAATAT	125037
983	983

1102 1102
 124086 CTCCTGTTTGTCTTATTATCAACACTTTCCTTCACATATATTGCT 124037
 1103 SerLysLeuTyrSerLeuPr 1109
 124036 CTCACGCTGTGTGCTTTACAGATTTGCCAACAAAGCTTCACACGGCACA 123987
 1109 oAsnProSerValLeuValSerLeuArgLeuAspPheLeuArgIleIleC 1126
 123986 GAATCCCACTCCCTGAATGCCCTGAGGATGACCTTTATCGCTATCGTGT 123937
 1126 ySerHisGlnHisTyrValThrLeuAsnLeuProCysSerLeuLeuThr 1142
 123936 GTAGCCATAGACGACGTACGTCATCTCAACCTGCCGTTCACCTCTCAGC 123887
 1143 ProProAlaSerProSerProSerValSerSerAlaThrSer.Gln... 1157
 123886 CCTCCAGCCTCGCCCTCACCTTCACCTCTCCACCACGCTCACAGGTGCT 123837
 1157 1157
 123836 ATTACATGCTGGAGGCTTAGACATTCATTACATTGCTCACAATATCC 123787
 1158 Se 1158
 123786 AATATCTTTTCTTCATTTCACCTTCTTATTATTTTACGTCCTCAGAG 123737
 1158 rSerGlyPheSerThrAsnValGlnAspGlnLysIleAlaAsnMetPheG 1175
 123736 TTCAGCATTTTCACGATGTGTCAAGATCAAGAGGCGCCACAATGTTCG 123687
 1175 LuLeuSerValProPheArgGlnGlnHisTyrLeuAlaGlyLeuValLeu 1191
 123686 AGCTCTCCGTCGCCGTTCGCCAATCAACATTTCTCTGCGCTCGCTTA 123637
 1192 ThrGluLeuAlaValIleLeuAspProAspAlaGlu.Gly..... 1204
 123636 TCAGACCTCTCTCATCTTGAATCTGTGATGTGAAAGGTAGAGACTGA 123587
 1205 LeuPheGly..... 1207
 123586 TCCTACTTCTTTTATATGTGTGCTGCTATTGGACCATGTCTAATGAT 123537
 1207 1207
 123536 ACTGAAATGCTGTATGCTTCAAGTCCAGTACATGAAAACTGAGTTGTT 123487
 1208 LeuHisLysLysValIleAsnM 1215
 123486 GGTTTTCTACACACAGCATTTTCTCTCCATATAAAAAAGCATTAGTG 123437
 1215 eValHisAsnLeuLeuSerSerHisAspSerAspProArgTyrSerAsp 1231
 123436 CTGTACACTCTCCGTGTGACGCCATGATGATTCGTGCAACAAGAC 123387
 1232 ProGlnIleLysAlaArgValAlaMetLeuTyrLeuProLeuIleGlyI 1248
 123386 CCTCAGGTAGGGCCCATGTGTCTCAGCTTACGTGCCCTCATTCACAT 123337
 1248 eIleMetGluThrValProGlnLeuTyrAsp..... 1258
 123336 TGTCAATGAGACACTGATACGCTTCATGACTTCTGTGCTAGTGTACAC 123287
 1259 PheThrGluThrHisAsnGln 1265
 123286 CCCCACACATTTTCATTAAGTTCGAAATGTTCACCTCGTGGCATTTTAAG 123237
 1266 Arg..... 1266
 123236 AATGTTAATTCATAGCTGTGTGTATGTTTGTGTAGACTCTCCCTG 123187

1267. GlyArgProIleCysIle..AlaThrAspAspTyrGluSerGluSerGly 1282
 123186 CTCGGGTCCGGGATGCTTCAGCCCATGACGATGCTGACCGATAGTGGC 123137
 1283 SerMetIleSerGlnThrValAlaMetAlaIleAlaGlyThrSerValPr 1299
 123136 AACACATCACTAGTCGTGTGGCCATGCAATTCGCCGCTCCCTTTAC 123087
 1299 O..... 1299
 123086 ACATGCCAAACCAACTCCTTTGCGCTACCCACAGCGGTAGATATATG 123037
 1300 GlnLeuThrArgProGlySerPhe..... 1307
 123036 GCTGGGGTCAGATGTTTCC..CCTGGCATTAACAGAAACAGGTACAGATC 122988
 1307 1307
 122987 CCACCTGTGTGGAGACATGTTCCTCAGCTTGACAGAACACACACTGG 122938
 1307 1307
 122937 CCTTCTGTAACTCCACTAAAAAACAACAAAATGTCATTTCTTCAC 122888
 1308 LeuLeuThrSerThrSerGly 1314
 122887 ATTGCTCATCTCCCTCAAGATTAAGCTGTGTGTGCTCATTAAGTGGC 122838
 1314 1314
 122837 TCTGTGCCCTCTCTAATAACTGTTCTTTCATGTTTGACATTTGAAGGGCA 122788
 1315 ArgGlnHisThrThr 1320
 122787 ACTTAGCGTGCTGTGTTGCTTTCGCCAGCCTGCCGCCACATCCAGCTCT 122738
 1320 heSerAlaGluSerSerArgSerLeuLeuIleCysLeuLeuTyrValLeu 1336
 122737 TGTCAAGAGATGTAGCCGGGACTGCTGGTGTGCTTCCTCGTGGGTGCTG 122688
 1337 LysAsnAlaAspGluThrValLeuGlnLysTrpPheThrAspLeuSerVa 1353
 122687 AAGATGACAGATCAGACCTCTCTGAGACCGCGGGGCTGTGATTTGTCAGT 122638
 1353 lleuGlnLeuAsnArgLeuLeuAspLeuLeuTyrLeuCysValSerCysP 1370
 122637 GCTGCMAATCAACCGCCTGCTGAGTCTGCTGACGCTGTGCATCTCGTGT 122588
 1370 heGluTyrLys..... 1373
 122587 TTGAATATAAAGTAAGTACTGTGTGATTTTGTCCCTTTGTTCACTGTC 122538
 1374 GlyLysLysVa 1377
 122537 TTGAGTCTGTGACATGTTTACTTTTGTATTATTAATAGGGGAAAAAGCG 122488
 1377 lPheGluArgMetAsnSerLeuThrPheLysLysSerLysAspMetArg 1394
 122487 TCTGGAGAGGATTAATAGCTGACATTAAAAATTCCTGACATCAAGG 122438
 1394 lAlaLysLeuGluGluAlaIleLeuGlySerIleGlyAlaArgGlnLeuMet 1410
 122437 CACGGTTGGAGAGGCCATCTAGTGGAAACATTTGGGGCTCGCTCAGAGATG 122388
 1411 ValArgArgSerArgGly..... 1416
 122387 GTTGTGCTGTGACAGAGTAATAAGTGTTCCTGCTGTGTGTGTGTG 122338
 1416 1416
 122337 TGCA 122288
 1417 Gln..LeuGluArgSerProSerGlySerAlaPh 1427

```

122287 TTCATGTTAATCTCAAAATCTTTTAGAGAGAGTCCC.....TA 122250
1427 eGlySerGlnGluAsnLeuArgTPrArgLysAspMetThrHisTPrArg 1444
122249 TAGCAGTCAAGGAATGTGAGGTGGAAGAAAGCAAGCTACTGCGGCC 122200
1444 lAsnThrGluLysLeuAspLys..... 1451
122199 AAAATACAGACAGAGTCGACCAAGTATGTTTTTTTAACTATTATG 122150
1451 ..... 1451
122149 CATTAATGCCAGTTACTGTAAAGATGTTCCATGTGCTATAGCACTG 122100
1451 ..... 1451
122099 AGTGTTCATCAGCGGCTGGGATACATCTGCTTTAATTTCCAG 122050
1452 .....Ser..A 1453
122049 TGCTTAATGCTGATGTGAAGTAAGAAATTAATCTTTATAGACTA 122000
1453 rGAlaGluIleGluHisGluAlaLeuIleAspGlyAsnLeuAlaThrGlu 1469
121999 AAGCTGATGAGAGCAGAGTCCGTGTGATGGAACCTTATCAACCGAG 121950
1470 AlaAsnLeuIleIleLeuAspThrLeuGluIleValGln..... 1483
121949 GCTTCTATATGCTCTTGACACGCTGAGATGTAGTCAGGTAGACAA 121900
1483 ..... 1483
121899 TGCTTCAAGCTGTGAAGTGCAGATATTGAGATTAATTTGCTT 121850
1484 .....ThrValSerValThrGluSerLysGlu 1492
121849 TCTTGTCTGTTGATTTTTCAGACTGTGTCATCAGACTGAAGA 121800
1492 uSerIleLeuGlyGlyValLeuLysValLeuLeuHisSerMetAlaCysA 1509
121799 AAGTGTCTGGGGAGATGATGAGTCTCTCCACAGCATGGCGAGAA 121750
1509 sncInserAlaValTyrLeuGlnHisCysPheAlaThrGlnArgAlaLeu 1525
121749 ACCAGAGTGCCTCTCTCTGACACTGCTTCACACAAAGAGCTTA 121700
1526 Val..... 1526
121699 GTTTTCAGGTACATAAATTAATTTTGTTAATGTGATGATATTT 121650
1526 ..... 1526
121649 CTCTTATTTCTTTTACAGTTGTCAGCATTTACATCATAGATTAGATC 121600
1527 .....Ser..Lys.... 1528
121599 ATACATACATATAGATAGAACATGTTGTGACCACTTGAACGTGAAGTGA 121550
1529 .....PhePro 1530
121549 AGGCAACTGAATACATGTGTATTATACCTGTCTGTAGTTCCCA 121500
1531 GluLeuLeuPheGluGluIuThrGluGlnCysAlaAspLeuCysLeuArg 1547
121499 GAGATGCTGTTTAAAGAGATACAGAACTTTGCGCAGCTATGTGTGG 121450
1547 GluLeuArgHisCysSerSerSerIleGlyThrIleArgSerHisPro 1564
121449 TCTTCTGCGTCACTGAGTAGCAGCGCTGCTGTCAAGAGTCAAGCGCT 121400
1564 eAlaSerLeuTyrLeuLeuMetArgGlnAsnPheGluIleGlyAsn... 1579
121399 CTGCTCTCTTTATCTGCTCATGAGACAGAACTTTGAGATTGGAATGTA 121350
1579 ..... 1579
121349 AGTAATATACTTTCATGTTGATGATGTAAATATACTTTCATTATGTTCA 121300
1579 ..... 1579
121299 ATACCTACTTAGGAAGTTTACTTGAGATTACTTAAAGAGTGACAAA 121250
1580 .....AsnPheAlaArg 1583
121249 GGAAGAGAAATTTACATCCCTTTATTTCTTTTGCAAACTTTGCTCGG 121200
1584 ValLysMetGlnValProMetSerLeuSerSerLeuValGlyThrSerGlu 1600
121199 GTAAAGATGCAAGTCAACCATGCTCATCTCATCTGTGGGAAACATCAAG 121150
1600 nAsnPheAsnGluLubPheLeuArgArgSerLeuLysThrIleLeuThrT 1617
121149 AAATTCATAGAGACATCTGCACACTGCTGTAAGAACTATCTTGACAT 121100
1617 TyrAlaGluGluAspLeuGluLeuArgGluThrThrPheProAspGln..V 1633
121099 ATGCTGAAGATGATGTGAACCTACGTGACACACCTTCCCGAGCAGGTA 121050
1633 al..... 1633
121049 TTAATGCGCCAGTGTAGAACATCCCATTTGTGAAGATGCATGAATGTTG 121000
1634 .....GlnAspLeuValPh 1638
120999 AGATTTAATTCATCTGTGTGTGAACACTGCTCCAGATTTGCTT 120950
1638 eAsnLeuHisMetIleLeuSerAspThrValLysMetLysGluHisGlnG 1655
120949 CAATTTGCATGATGTCTCACCGATACGTGTAATGAAGAAACACAGC 120900
1655 lAspProGluMetLeuIleAspLeuMetTyrArg..Ile..... 1667
120899 AGGATCCAGAGATGCTTATTACCTGATGTACAGTATATGCGATGCAT 120850
1667 ..... 1667
120849 TAATATCAAAATATAGTATCAGACTTAATATCTGCTTTTGTAAATC 120800
1668 .....AlaIysGlyTyrGlnThrSerProGlu..ArgLeuThrTrpLe 1681
120799 AGGATTTGCAAGGGCTATCAAAACCTCTCCGACCTGCTGTGACTTGCT 120750
1681 uGlnAsnMetAlaGlyLysHisSerGluArgSerAsnHisAlaGluAla 1698
120749 CCAGAACATGGCAGAAACACCTCTGAGAGGAGAAACCATCTGTAACCTG 120700
1698 lAgGlnCysLeuValHisSerAlaAlaLeuValAlaGluTyrLeuSerMet 1714
120699 CCCACTGCTGCTGTCACAGTGCAGACTGTGTCAGAAATACCTCAACATG 120650
1715 LeuGluAspArgLysTyrLeuProValGlyCysValThr..... 1727
120649 CTGAGGAGATGTGCTACTTGCATCGTGGTGGCTGACCTTTACAGTGG 120600
1727 ..... 1727
120599 TGAGAGACGCTGACTTTAAGTATGTATGCAGCAGCTAGTTGGCTGGTG 120550
1728 .....PheGlnAsnIleSerSerAsnVal 1735
120549 TTATAAATGCATCATACATACCTTTCAAGAACATTTTCATCAACCTC 120500
1736 LeuGluGluSerAlaValSerAspAspValValSerProAspGluGlu 1752
120499 CTGAGAGAGTCTGTGTGTGATGACGTTGTCTCCGAGAGAAAGGG 120450

```



```
118649 AAAACAACACACCTGCTGTCATTTCTCAGAGGTGTAG 118650
|||||
2064 AspAlaLeuArgLysAnlysSerLeuIleGlyProValGlnLysGluTy 2080
|||||
118599 GATGCCCTGAGAGAAATTAAGCCCTCATTTGGCGGATCAGAGAGATA 118550
|||||
2080 rGlnArgGluLeu 2084
|||||
118549 CCACAGAGAGCTG 118537
seq_name: gb_hlg:AC017375

seq_documentation block:
LOCUS AC017375 73515 bp DNA linear HTG 09-DEC-1999
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, In ordered
pieces.
ACCESSION AC017375 GI:6553611
VERSION AC017375.1
KEYWORDS HTG: HTGS_PHA5E2.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 73515)
REFERENCE 1 Adams,M. and Venter,J.C.
AUTHORS Direct Submission
TITLE Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive,
JOURNAL Rockville, MD, USA
COMMENT This sequence was identified as CDM:10210477 by the submitter.
For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
location/Qualifiers
source 1..73515
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
BASE COUNT 21627 a 14907 c 14713 g 22268 t
ORIGIN
```

```
alignment_scores:
Quality: 4700.00 Length: 2339
Ratio: 2.892 Gaps: 57
Percent Similarity: 69.474 Percent Identity: 44.164

alignment_block:
US-09-737-246-2 x AC017375/rev ..
Align seg 1/1 to reverse of: AC017375 from: 1 to: 73515

4 ArgArgLaphealagInlysIleSer..... 12
:::|||||:|||||:|||||:|||||:|||||
43211 CAGGCGCTGTTGTGCAAAAAGTTTCCAAAGTAAGTTTTCATAAGGAA 43162
12 ..... 12
43161 ATAGTTATTTTAAACAAGAACGCCCTTAATGCTGTATCATATATGTA 43112
13 ..... ArgThrValAlaIleGluValArgLys 21
|||||
43111 CATATATCATTTTGTAAATCCCTAAGCAGTCGGCTGCGGATGTACGAAA 43062
22 GlnIleSerGlyInTySerGlySerProGlnLeuLeuLysAnlysLeuAs 38
:::|||||:|||||:|||||:|||||:|||||
43061 AATGTTCTGCTGTCATATCAAAAGCAATG.....GACCCATC 43021
38 nIleValGlyAnlysIleSer..HisHisThrThrValProLeuThr..... 52
:::|||||:|||||:|||||:|||||:|||||
43020 ATTGTGTGTTCTTCGGTGAAGACATATTGTTAAATATACATTTT 42971

53 .....GluAlaValAspProValAs 59
:::|||||:|||||:|||||:|||||:|||||
42970 CTGTTAAGATTCATTTTTCATATAGATCTCCCGCAGACAGCCCTTGA 42921
59 pLeuGlnAspTryLeuIleThrHisProLeuAlaValAspSerGlyProL 76
|||||:|||||:|||||:|||||:|||||
42920 TTACGAAGAAATTTCTATCCCGCATATGAATCATCAACCGGGATCCCT 42871
76 euArgAspLeuIleGluPheProAspAspIleGluValValTySer 92
|||||:|||||:|||||:|||||:|||||
42870 TGAACAATATATTGACTTCTCAAGGGATGTAACTGTTAAATCATTT 42821
93 ProArgAspCysArgThr..LeuValSerAlaValProGlu..... 105
|||||:|||||:|||||:|||||:|||||
42820 CCCCCAAATACAGACATGTTGATCAGCTGTTCCTCCATAGAAATTGTA 42771
106 .....Glu 106
42770 TGTATGCACCTAAGTTATACAACAGCTACATAGCTACAGTCTTTTAGG 42721
107 SerGluMetAspProHisValArgAspCysIleArgSerTryThrGluAs 123
|||||:|||||:|||||:|||||:|||||
42720 TCTGATTTGCCCATACAGCTACAGAAATGTGTAACTGCTATAGCGGCC 42671
123 pThrAlaIleVal.....I 128
42670 ATGGAAGTGTGTAATATGCACAGCGCACCTTCCAGTCTTGCTATA 42621
128 leArgLysTryHisLysLeuGlyThrGlyPheAsnProAsnThrLeuAsp 144
|||||:|||||:|||||:|||||:|||||
42620 TCCTGGACAGCTATCGACCGCGGAAC..ATAAGTCCCTGCGCTACAC 42574
145 LysGlnLysGluArgGlnLysGlyLeuProLysGlnValPheGluSerAs 161
:::|||||:|||||:|||||:|||||:|||||
42573 CAGGAATTCGAGATCGACAAAGAT.....TTTGTGTCATT 42539
161 pGluAlaProAspGlnAsnSerTyGlnAspAspGln..AspAspLeu 177
|||||:|||||:|||||:|||||:|||||
42538 TGA.....GAGCGCTTCGCTTATAAAGCAGAGCTGCACCTCAAGTT 42495
177 ysArgArgSerMet.....SerIleAspAsp... 185
:::|||||:|||||:|||||:|||||:|||||
42494 CTCGACAAATCCATTGCAAGTTGGCTCAGTAAGCTCTTACGCGACACT 42445
186 ...ThrProArgLysSerThrAlaCysSerIlePheAspLeuLysAsnSe 201
|||||:|||||:|||||:|||||:|||||
42444 TTGACCCCGCGTGTCTTGGGCTAGC.....TTTGATCTTCGACGCTC 42401
201 rLeuProAspAlaLeuLeuProAsnLeuLeuAspArgThrProAsnGlu 218
:::|||||:|||||:|||||:|||||:|||||
42400 GGTATATGACCCCATATATACCAATTTGCTTGACATATGACTCCGGAGC 42351
218 lIleAspArgGlnAsnAspAspGlnArgLysSerAsnArgHisLysGlu 234
:::|||||:|||||:|||||:|||||:|||||
42350 AATTTGATCATCAATCAATGAAGCAGCCGCGCAGCAAGATCGCAATGAGCT 42301
235 LeuPheAlaLeuHisProSerProAspGlnGluProIleGluArgLe 251
|||||:|||||:|||||:|||||:|||||
42300 CTTTCTCAGCTATCCGGAATCTGAAGCGGAGGAAACATTGAAGCGCG 42251
251 uSerValProAspIleProLysGlnHisPheGlyGlnArgLeuLeuVal 268
:::|||||:|||||:|||||:|||||:|||||
42250 CCGTCCGCGAGAGATTCAGTTGAACATATGGCTCATTCAGTTAGTTA 42201
268 ysCysLeuSerLeuLysPheGluIleGluIleGluProIlePheAlaSer 284
:::|||||:|||||:|||||:|||||:|||||
42200 ACTGCTTACAGCTCGGCTTGAAGCTGGAAGTGGACCAATTTTTCATTCG 42151
285 LeuAlaLeuTyAspValLysGlnLysLysLysIleSerGluAsnPheTy 301
:::|||||:|||||:|||||:|||||:|||||
42150 ATGCTATTATATGATGCTAAGACGCGCAAAATTTCTGAATAATTCTTA 42101
```

[illegible]

40353ACCAACAGCGACTAATCATGCAATGCGCA 40326
 882 eAlaValaIarProAlaSerLeuAsnLeuAsnArgSerArgSerLeuSer 898
 ||||| |||||
 40325 GCACACAGAA.....GACTTGCAATTAGCAATCGTACACCAATATAT 40282
 899 AsnSerAsnProAspIleSerGlyThrProThrSerProAspAspGluVa 915
 :|||:
 40281 GATTAAT..... 40276
 915 IarSerIleIleGlySerLysGlyLeuAspArgSerAsnSerTrpVala 932
 :|||:
 40275 ...GTTTACGACAGTGGCCGACGCTGATCGCAAGGTAA.....A 40237
 932 snThrgIyGlyProLysAlaIaIaProItpGlySerAsnProSerProser 948
 |||||
 40236 ATACATATTGTTGTAAGCT..... 40217
 949 AlaGluSerThrGlnAlaMetAspArgSerGlyAsnArgMetSerSerHi 965
 |||
 40216TTAATATACCTTTTCT 40197
 965 sThrGluThrSer...SerPheLeuGlnThrLeuThrGlyArgLeuProT 981
 :|||
 40196 TTCAGAGCTGCTCATAGATGTCTTCACCTCGATGGCTGACGGGATGCC 40147
 981 hLysLysLeuPheHISGluGluLeuAlaLeuGlnTrpValaLysSer 997
 :|||:
 40146 AAGTGGGCTGCTTCACGAGAAATTAGCTTCATGGGTGTTGCTCAAGT 40097
 998 GlySerValArgLysLeuSerAlaLeuGlnGlnAlaTrpPhePheGluLe 1014
 |||||
 40096 GGAAAGGCGACGCTAGCTTACCATGCAATGCGTGTATTATTTAGACT 40047
 1014 uMetValLysSerMetValHISLeuThrPheAsnAspLysLeuGluA 1031
 :|||:
 40046 TATGTCAATCTATCATTTGAACATTGTCATTTGCAATCTTAAATG 39997
 1031 LaProArgLysSerArgPheProGluArgPheMetAspIleLeuAla 1047
 :|||:
 39996 GTCCCGGCAACATCGATTCCACATTCAGTTTAATGATGATCTATCCAC 39947
 1048 LeuValSerThrIleAlaSerAspIleValSerArgPheGlnLysAspTh 1064
 |||||
 39946 CTGGTGCACCTTGATTACACAAAGAGTGGCGGTACACAGACAGCAAC 39897
 1064 rGluMetValGluArgLeuAsnThrSerLeuAlaPhePheLeuAsnAspL 1081
 :|||:
 39896 TAAATTAGCGCGATCATTTAAATGCAAGTTTATGCTTTTATATTGACA 39847
 1081 euLeuSerValMetAspArgGlyPheValPheSerLeuIleLysSerGys 1097
 :|||:
 39846 TACTTAAGCATCATGATCGGGATTTGTTTGCGCTCATCAAAACGTAC 39797
 1098 TrpLysGlnValSerSerLysLeuTrpSerLeuProAsnProSerValIle 1114
 |||
 39796 ACAAGGTTCTGATTTCTAAATAATGCTTCATACCGGAT.....CT 39756
 1114 uValSerLeuArgLeuAspPheLeuArgIleIleCysSerHISGluHIS 1131
 :|||:
 39755 AATGAACTATTAATAATGACTTTTGTGAAATGTGTGTACTCATGAGCATT 39706
 1131 TrpAlaThrLeuAsnLeuPro.....CysSerLeuLeuThrPro 1143
 :|||:
 39705 TTGTAGCTTAACCTTGCCGTTTGTACTTCATACCAATGGTAACAGCG 39656
 1144 ProAlaSerProSerProSerValSerSerAlaThrSerGlnSerSerG 1160
 |||||
 39655 CCATGACGCCGCCACGCAAGCAACAGCTGACAGCAACGCCAACTCTTGG 39606
 1160 YPheSerThrAsnVal..... 1165
 |||
 39605 CGTAAGTAAAAAACAATACACTTTAATTTTACTTATTTGTGGCTTCTT 39556

1166GlnAspGln..LysIleAlaAsnMetPheGluLeuSerValPro 1180
 :|||:
 39555 TTAAAGGATCATAGACAGACCCCTTCACGCCGACCTTAGTCAAGAT 39506
 1180 heArgGlnGlnHISTrpLeuAlaGlyLeuValLeuThrGluLeuAlaVal 1196
 |||
 39505 TCCTCCAGCAACATTTTCTAGTCGGACTTGTCTGAGCAGCACTTAGCACA 39456
 1197 IleLeuAsp...ProAspAlaGluGlyLeuPheGlyLeuHISLysLysVa 1212
 :|||:
 39455 GTAATGAGGTGCCAAACCCCA.....TTACATGGAAGAAC 39418
 1212 ILeAsnMetValHISAsnLeuLeuSerSerHISAspSerAspProArgT 1229
 :|||:
 39417 AATAAGCTGATTTCGCACACCTGATGACATCTCAGATTTTGACGCCGCG 39368
 1229 TrpSerAspProGlnIleLysAlaArgValAlaMetLeuTrpLeuProLeu 1245
 |||||
 39367 ATACGAAACCTGATGCACGTGCAAGGCTGCTGTTATATATACCGCTA 39318
 1246 IleGlyIleIleMetGluThrValProGlnLeuTrpAspPhe..... 1259
 :|||:
 39317 CTTTCTATAGTATGATGATGCATTCCTCAGTTGCATCAGCAGCGCCTTGA 39268
 1260 ThrGluThrHISAsnGlnArgGlyArgProIleCysIleAlaThrAspA 1276
 :|||:
 39267 GCAGATCGTCTGCACCAATCGTCA.....CTTGAG 39233
 1276 sPTrpGluSerGlu.....SerGlySerMetIleSerGlnThr 1288
 |||||
 39232 ACTATCAAGGCCACATCAAACTATTATACAGATCAACCAATTCGCGAA 39183
 1289 ValAlaMetAlaIleAlaGlyThrSerValProGlnLeuThrArgProG 1305
 |||||
 39182 GTGCAATTGCGATATGCGGA.....AGTGTCCCTA 39151
 1305 YSerPheLeuLeuThrSerThrSerGlyArgGlnHISThrTrpPheSerA 1322
 :|||:
 39150 CTCCTACTTG.....AATGACAAGTTAAACAAATTCGCCCTCGACT 39107
 1322 IarGluSerSerArgSerLeuLeuIleCysLeuLeuTrpValLeuLysAsn 1338
 :|||:
 39106 CGGAAATATCCCGCATTTATGTTGTTGTTTCTGTCGGTGTAAACAAAC 39057
 1339 AlaAspGluThrValLeuGlnLysTrpPheThrAspLeuSerValLeuG 1355
 :|||:
 39056 TTGAGCGCAACGTTTATATATGCTGCTTATGATTAAGATCCGCAATG 39007
 1355 nLeuAsnArgLeuLeuAspLeuLeuTrpLysCysValSerCysPheGluT 1372
 :|||:
 39006 AGTCACCAAAATGCTTCAAGGTGTTAACGTTTCTGTTAAAACTTTCCAGT 38957
 1372 TrpLysGlyLysLys.....ValPheGluArgMetAsnSerLeuThr 1385
 |||
 38956 ACACAGCAAAAAAAGCGTGGCCACCTTAAACGAAATACCAAAAT 38907
 1386 PheLysLys.....SerLysAspMetArgAlaLysLeuGlnGluAlaI 1400
 |||||
 38906 TTTAGGAAACATGATCTACTGATGTTAAGGAAAGCGTGAAGATATAT 38857
 1400 eLeuGlySerIleGlyIaArgGlnGluMetValaIarGlySerArgLys 1417
 |||
 38856 TAGAGCACAAAATTCAGACCTTATGATTTAATAATCGTCGAAAAAGATC 38807
 1417 nLeuGluArgSerProSerGlySerAlaPheGlySerGlnGluAsnLeu 1433
 :|||:
 38806 GT.....AATCAACAGAAATAATTC 38787
 1434 ArgTrpArgLysAspMetThrHISTrpArg...GlnAsnThrGluLysLe 1449
 |||||
 38786 CGATGAGGAAAGACAGATGATCATATGCTCAATATATGCGGATGCTGT 38737

36942 TCGCATATTGTAATTAAGCGACAGCAAAATAGTCTGGAAACCGATTGAAG 36893
1985 AAlAlAlleglu..... 1988
36892 TGGCATTTGAAGTAAGTATGTGTAAATCCCAAAAGTTATAATAATTA 36843
1989AspMetGlnLysLysThrgIngluLeuAlaPheAl 2000
36842 AACGTTATTGGCTAGGATATTCAAAAAACAATTGGAAATGGACAGCTGC 36793
2000 aThHisGlnAspProAlaAspProLysMetLeuGlnMetValLeuGlnG 2017
36792 CACGATCAAGACGACGACGCCCAAAATTTTCAAATGGTACTGAGG 36743
2017 1ySerValG1yThrThrValAsnGlnGlyProLeuGluValAlaG1yVal 2023
36742 GATGATTGGAAACCACTGTTAACCAAGACCGATGCAATATGGCAGTGTG 36693
2034 PheLeuSerGluLeuProSerAspProLysLeu...PheArgHisHisAs 2049
36692 TTCCCTTCCATTATCCGACGAAACAACGTACCGACAAACACCAAAA 36643
2049 nLysLeuArgLeuGlySerPheLysAspPheThrLysArgCysGluAspAlaL 2066
36642 CAAACTTGGTTGGCTTCCGCGAGGTTTTCAAAGCGTTGCTGATGCTT 36593
2066 euArgLysAsnLysSerLeuLeuGlyProValGlnLysGluTyrGlnArg 2082
36592 TGAATAAGATCGCAATCTACTTTCAGATCAAAAAAGATTACCAACGA 36543
2083 GlnLeu 2084
36542 GAACGTG 36537
seq_name: gb_in:AC008318
seq_documentation_block:
LOCUS AC008318 166626 bp DNA linear INV 05-APR-2001
DEFINITION Drosophila melanogaster, chromosome 2L, region 21A-21A, BAC clone
BACR09J20, complete sequence.
AC008318
AC008318.7 GI:13549312
VERSION AC008318.7
KEYWORDS fruit fly.
SOURCE Drosophila melanogaster
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS
1 (bases 1 to 166626)
Celisner,S.E., Adams,M.D., Krommiller,B., Tyler,D., Wan,K.H.,
Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,
Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
Dodson,K., Dorset,L.V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
Farrleria,S., Frise,E., Galle,R.F., Gary,N.S., George,R.A.,
Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
Idegami,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
Phouaneavong,S., Piltman,G.S., Puri,V., Richards,S., Scheeler,F.,
Stapleton,M., Strong,R., Svirkas,R., Tector,C., Williams,S.M.,
Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
Sequencing of Drosophila chromosome 2L, region 21A-21A
Unpublished
2 (bases 1 to 166626)
Celisner,S.E., Agbayan,A., Arcina,T.T., Baxter,E., Blazej,R.G.,
Butenoff,C., Champe,M., Chavez,C., Chew,M., Cislolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karia,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
Pfeiffer,B., Poon,U., Sequeira,A., Sethi,H., Snir,E.,
Svirkas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and

Rubin,G.M.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT On Apr 5, 2001 this sequence version replaced gi:6728945.
Sequence submitted by:
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
This sequence was assembled using end sequences from a whole genome
shotgun and from subclones of this BAC and its neighboring clones.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive web site (<http://www.fruitfly.org/sequence/>) or send email
to bdg@fruitfly.berkeley.edu.
FEATURES
source
1.166626
/organism="Drosophila melanogaster"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="2L"
/map="21A-21A"
/clone="BACR09J20 (D841)"
/clone_lib="RPCT-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBACE3.6)"
BASE COUNT 49737 a 32951 c 34017 g 49921 t
ORIGIN
alignment_scores:
Quality: 4690.00 Length: 2340
Ratio: 2.886 Gaps: 58
Percent Similarity: 69.444 Percent Identity: 44.188
alignment_block:
US-09-737-246-2 x AC008318 ..
Align seg 1/1 to: AC008318 from: 1 to: 166626
4 ArgArgAlaPheAlaGlnLysLieser..... 12
:::||||:||||:||||:||||:||||: 12
85506 CACGCGCTGTTGTGCAAAAAGTTCCAAAGTAGTTTTCCTAAAGAA 85555
12 12
85556 ATAGTTATTTTAAACAAGAAAGCCCTTAATGCTGTATACATATGTA 85605
13ArgThrValAlaAlaGluValArgLys 21
||| |||||:||||:||||: 21
85606 CATATATCAATTTGTAAATCCCTAGGACGTCGGCTGGATGACGAAA 85655
22 GlnLieserGlyGlnTyrSerGlySerProGlnLeuLeuLysAsnLeuAs 38
:::|||||:||||:||||:||||: 38
85656 AATGTTCTGCTGTCCTATCATCAAGGCAATG.....GACCCATC 85696
38 nLysValG1yAsnLieser HisHisThrThrValProLeuThr..... 52
::: ||||:||||: 52
85697 ATTGTGTGTTCTTCGGTAGAGCACATGTTTAAATATFACAAATTTT 85746
53GluAlaValAspProValAs 59
::: ||||:||||: 59
85747 CTGTTAAGATTCATTTTTCATATAGATCTCCCGACAGCCCTTGA 85796
59 pLysGluAspTyrLeuLeuThrHisProLeuAlaValAspSerGlyProL 76
|||:||||:||||: 76
85797 TTACGAAGAAATTTCTATCCAGCATATGAACATCAACACGGGATCCCT 85846
76 euArgAspLeuLeuGluPheProProAspLysLiesGluValLysSer 92
|||:||||:||||: 92
85847 TGAACATATATTTGACTTTCCTCAAGGAGATGATCTGTTAAATCATTT 85896
93 ProArgAspCysArgThr.LeuValSerAlaValProGlu..... 105

635 ysluProAlatThrLeuThrAspHisHisLeuLeuPheThrPheThr 651
 ||||| : : : : :
 87693 CGCTGCTGCATCAATAAGACAGATCACCATTATTGTCACCATTTAC 87742
 652 HisValSerCysGlnGlnLeu.....AsnThrProLeuGluTh 665
 ||||| : : : : :
 87743 CATGTTTCATGCAAAAGAACACAGGACCTTGACCGCTGAGAGAC 87792
 665 rProAlaGlyThrThrTrpIlePrometLeuGlnAsnGlyArgLeuYST 682
 ||||| : : : : :
 87793 TCCTATGCGCTACACATGTTGCTTACTGAGAGATGAGAAACTTAAGT 87842
 682 hGlnGlnPheCysLeuProAlaSerLeuGlnGlyProGlnAlaLys 698
 ||||| : : : : :
 87843 TTGGGAAATTAAATCTCTGTTATGTAGAAATCACCGCGAAATTTAC 87992
 699 SerValLeuSerProGluValProLeuProGlyMetLysTrpValAsp 715
 ||||| : : : : :
 87893 TCATTATATACACCGAATGTTCCACTGCTGGAATTAATGCTGACAA 87942
 715 hHisLysGlyValPheAsnValGluValAlaValSerIleHisT 732
 ||||| : : : : :
 87943 TCACAGAGCGCTGTTCTATTATATGTAAGACAGTCAACCGCAATTCATA 87992
 732 hGlnAspProTyrLeuAspLysPhePheAlaLeuValAsnAlaLeuAsp 748
 ||||| : : : : :
 87993 CTTTGAGATTCTTCTCGATCATCTCTTAATTTGGCAATATCTTAC 88042
 749 GluHisLeuPheProValArgIleGlyAspMetArgIleMetGluAsn 765
 ||||| : : : : :
 88043 ACTCGAAATATACCTTCCCATATTTGC.....GAAATATA 88077
 765 nLeuGluAsnGluLeuLysSerSerIleSerAlaLeuAsnSerSerGln 782
 ||||| : : : : :
 88078 TATGAAAGCAATTAAGAAATGCTGCTGATATCGAATATGCAAAATC 88127
 782 euGluProValAlaArgPheLeuHisLeuLeuAspLysLeuLeu 798
 ||||| : : : : :
 88128 GTGAGCCATTACTGAGGCATCTACCATTAAGTCTTGATTAAGTTTCGA 88177
 799 LeuValIleArgProProValIleAlaGlyGlnIleValAsnLeuGly 815
 ||||| : : : : :
 88178 TTGTTAGTTGTCAACCCACACAAAGTGTGAGCAACCAATGCTCTGAGATC 88227
 815 nAlaSerPheGluAlaMetAlaSerIleLeuAsnArgLeuHisLysAsn 832
 ||||| : : : : :
 88228 TACGGTTTTCGAAGTTCTTGTGTATCATCTTGTATCATATATCTTA 88277
 832 euGluGlyAsnHisAspGlnHisGlyArgAsnSerLeuLeuAlaSerTyr 848
 ||||| : : : : :
 88278 ACGACGATCAATATGACCAATACGGTGGCAAAAGTTGCTATACACATAT 88327
 849 IleHisTyrValPheArgLeuProAsnThrTyrProAsnSerSerPr 865
 ||||| : : : : :
 88328 GTGCACTTTCATATGTAATAATTCGACACCCCTTCA..... 88363
 865 oGlyProGlyGlyLeuGlySerValHisTyrAlaThrMetAlaArgS 882
 ||||| : : : : :
 88364AGCAACACGCGACTAACATCGACGCA 88391
 882 eAlaValArgProAlaSerLeuAsnLeuAsnArgSerArgSerLeuSer 898
 ||||| : : : : :
 88392 GCACACACAGAA.....GACTTGCATTTAAGCAATCTGACACCATATAT 88435
 899 AsnSerAsnProAspIleSerGlyThrProThrSerProAspAspGluVal 915
 ||||| : : : : :
 88436 GATAT..... 88441
 915 lArgSerIleIleGlySerLysGlyLeuAspArgSerAsnSerTrpVal 932
 ||||| : : : : :
 88442GTTTATGCAAGTGGCGAAGCTTGATGCAAAAGTAA.....A 88480

932 snThrGlyGlyProLysAlaAlaProTrpGlySerAsnProSerProSer 948
 ||||| : : : : :
 88481 ATACATATTTTGTAAAGCT..... 88500
 949 AlaGluSerThrGlnAlaMetAspArgSerCysAsnArgMetSerSerH 965
 ||||| : : : : :
 88501TTAAATATACCTTTTTCCT 88520
 965 sThrGluThrSer...SerPheLeuGlnThrLeuThrGlyArgLeuPro 981
 ||||| : : : : :
 88521 TTCAGACCTTCATAGATGCTCTTCACTCGATGCGTGCAGGAGATGTC 88570
 981 hLysLysLeuPheHisGlnGluLeuAlaLeuGlnThrTrpValValCysSer 997
 ||||| : : : : :
 88571 AAGTGGCGCTCTTCAAGAAATTAAGCTTGATGGGTGTTGCAAGT 88620
 998 GlySerValArgGluSerAlaLeuGlnGlnAlaThrPhePheGlu 1014
 ||||| : : : : :
 88621 GGAAGGCGACTGACTTACGATATCAAAATTCGTGTTTATTTAGACT 88670
 1014 uMetValLysSerMetValHisHisLeuTyrPheAsnAspLysLeuAla 1031
 ||||| : : : : :
 88671 TATTTGCAATCTATGATTTGACATTTGCAATTTGCAATATCTTAATG 88720
 1031 lAProArgLysSerArgPheProGluArgPheMetAspIleAlaAla 1047
 ||||| : : : : :
 88721 GTCCCGCAAAATCATGATTTCCACATCATGTTTAAATGATATATATCACC 88770
 1048 LeuValSerThrIleAlaSerAspIleValSerArgPheGlnLysAsp 1064
 ||||| : : : : :
 88771 CTGTTGCACTTATGATCAACAAAGGTGCTGCTACACAGCAACAAAC 88820
 1064 rGluMetValGluArgLeuAsnThrSerLeuAlaPhePheLeuAsnAsp 1081
 ||||| : : : : :
 88821 TAAATTTAGCGCAGTCAATTAATGCAAGTTTATTTATTTTGACA 88870
 1081 euLeuSerValMetAspArgGlyPheValPheSerLeuIleLysSerCys 1097
 ||||| : : : : :
 88871 TACTAAGCATCATGATGATGGGATTTGTTTCCGCCCAAAAGCTAC 88920
 1098 TyrLysGlnValSerSerLysLeuTyrSerLeuProAsnProSerVal 1114
 ||||| : : : : :
 88921 ACAAAAGTTTCGATTTTAAAAATGCTTCAATACCGAT.....CT 88961
 1114 uValSerLeuArgLeuAspPheLeuArgIleIleCysSerHisGluHisT 1131
 ||||| : : : : :
 88962 AATGAACATTAATTAACATTTTGAATTTGTGTAGATGATGAGAT 89011
 1131 yValThrLeuAsnLeuPro.....CysSerLeuLeuThrPro 1143
 ||||| : : : : :
 89012 TTGTAGCTTAACTGCGCTTGTGTTACTTCATACCAATGTAAACAGCG 89061
 1144 ProAlaSerProSerProSerValSerSerAlaThrSerLysSerGln 1160
 ||||| : : : : :
 89062 CCAGTCAAGCCCGACGCAAGCAACGTGACAGCAACGCAAACTTCTTG 89111
 1160 yPheSerThrAsnVal..... 1165
 ||||| : : : : :
 89112 CGTAAGTAAATAAACATATACCTTTATTTTACTTATTTGTGGCTCTT 89161
 1166GlnAspGln...LysIleAlaAsnMetPheGluLeuSerValProp 1180
 ||||| : : : : :
 89162 TTAAGAGGATCACTAGAGAGCCCTTCAACGCGCACTTATGCAAGAT 89211
 1180 hArgGlnGlnHisTyrLeuAlaGlyLeuValLeuThrGlnLeuAlaVal 1196
 ||||| : : : : :
 89212 TCTGTCAGCAACATTTCTAGTGGACTTGTCTGAGACGACTTATGAGCA 89261
 1197 lLeuAsp...ProAspAlaGlyLeuPheGlyLeuHisLysLysVal 1212
 ||||| : : : : :
 89262 GTATTTGAGGTGCCAAACCCCA.....TTACATGGAATAAGC 89299
 1212 lIleAsnMetValHisAsnLeuLeuSerSerHisAspSerAspProArgT 1229

93900	AAATAAGGCTATTCGCAACGTCGATGACATCTCACGAGTATTGGACCGCGCT	89349
1229	yrSerAspProGlnIleLeuAlaArgValAlaMetLeuTyrLeuProLeu	1245
89350	ATAGGGAACATGATACAGTCGCAAGGGGCGCTCTTATATATATACCGCTA	89399
1246	lleGlylleIleMetGluThrValProGlnLeuTyrAspPhe.....	1259
89400	CTTTCTATAGATGATGATTCGATTCCTCAGTTGCATCAGCAGCGCTTGA	89449
1260	.ThrGluThrHisAsnGlnArgGlyArgProIleCysIleAlaThrAspa	1276
89450	GCAGATCGCTCGCAGCAATCGGTCAA.....CTTGAAC	89484
1276	sPrryGluSerGlu.....serGlySerMetIleSerGlnThr	1288
89485	ACTATCAAGGGCCACATCAACATATATACGATCAACCATTAATCCGGAA	89534
1289	ValAlaMetAlaIleAlaGlyThrSerValProGlnLeuThrArgProG	1305
89535	GTGGCATTTGGCATATCCGGA.....AGTCGTCCTTA	89566
1305	YserPheLeuLeuThrSerThrSerGlyArgGlnHisThrThPheSera	1322
89567	CTCCACTGTG.....AATGAACAAGTTAAACAAATTCGCCCTCAGCT	89610
1322	IaGluSerSerArgSerLeuLeuIleCysLeuLeuThrValLeuYAsn	1338
89611	CGGAAATACCCGCCCATTTATGTGTGTCTTCTGTGGTCTTAATAAAC	89660
1339	AlaAspGluThrValLeuGlnLysTrpPheThrAspLeuSerValLeuG	1355
89661	TTGGAGCCCAACGTTTATATCTCGGCTTATGAGTTAAGTCCGCAATAG	89710
1355	nLeuAsnArgLeuLeuAspLeuLeuTyrLeuCysValSerCysPheGluT	1372
89711	AGTGCACCAATAGCTTCAAGGTGTTAACGTTTCTGTAACAACTTTGCAAG	89760
1372	YrLysGlyLysLys.....ValPheGluIleArgMetAsnSerLeuThr	1385
89761	ACACAGGACAAAAAAGCGTGCCACCTTAACAGCAAGATATACAAAGT	89810
1386	PheLysLys.....SerLysAspMetArgAlaLysLeuGlnLualI	1400
89811	TTTAGAAACATCGATCTACTAGATGTTAAGAAAGCTGGAGAAATGTAT	89860
1400	eLeuGlySerIleGlyAlaArgGlnLysMetValArgArgSerArgGlyG	1417
89861	TAGAGGCCCAAAATTCACGACGTTATGATTTAATTCGTCGAAAAAGATC	89910
1417	IleuGluArgSerProSerGlySerAlaPheGlySerGlnLysAsnLeu	1433
89911	GT.....AATTCAACAGAAAAATTC	89930
1434	ArgTrpArgLysAspMetThrHisTrpArg...GlnAsnThrGluLysLe	1449
89931	CGATGGAGGAAGACACGATGCCATATCGCTCTCAATATAGCGGATCGCT	89980
1449	uAspLysSerArgAlaLysIleGlyHisGlyAlaLeuIleAspGlyAsnL	1466
89981	GGGAAAAAGTGACATGACTGAATTAATCACTTCATTCATCGAAGCGCTTT	90030
1466	euaIaThrGluAlaAsnLeuIleIleLeuAspThrLeuGlnLysValAl	1482
90031	TGGCACTGAAGTTGCCACTGTACTTCTCGATACCCCTTAATAATATGTT	90080
1483	GlnThrValSerValThrGluSerLysGluSerIleLeuGlyGlyValLe	1499
90081	CAT.....GCTGCTGCCAACCTTTATACAAATCTCTTGGAACGTGCT	90124
1499	uLysValLeuLeuHisSerMetLacYsaAsnGlnSerAlaValTyrLeuG	1516
90125	AAAGTGCTTCTTCATCTTTGTGCCGAATCAATCGTTTGGCCCTGC	90174
1516	IleHisCysPheAlaThrGlnArgAlaLeuValSerLysPheProGluLeu	1532
90175	AAACTGTGTCCTCCCAAGCGCGTTAATCTTAAATTTCCAAACTTA	90224
1533	LeuPheGluGlnGluThrGlnGlnCysAlaAspLeuCysLeuArgLeuLe	1549
90225	TTATTTGATGACGAGACCGAATATATGCGCGAATTTGTGCTATTTTACT	90274
1549	uArgHisCysSerSerSerIleGlyThrIleArgSerHisProSerAlas	1566
90275	GAACATTTGGGGCCCTATTCGAGGAATATGATCAGCAGCGATGCAT	90324
1566	eLeuTyrLeuLeuMetArgGlnAspPheGluIleGlyAsn.....	1579
90325	CACATATTTACTATATAGACAAAAATTTGAAATGTGAATGTAACTTT	90374
1579	1579
90375	ATTTTATGTTTACGTTTATTAACTAATAATATGTCATATTCGTT	90424
1580	.Asn.PheAlaArgValLysMetClnValProMetSerLeuSerSerLeu	1595
90425	AGAATTTGGCCCGGATTAAATGATCAAGTGCAGATGCTTTAAGTCCTTA	90474
1596	ValGluThrSerGlnAsnPheAsnGlnGluPheLeuArgArgSerLeuL	1612
90475	GTTGGAACAAGTTCGGCTTTATGTAGCAATCTTAAGCCGTCGCTTAA	90524
1612	sThrIleLeuThrTyrAlaGlnLysAspLeuGluLeuArgGluThrThrP	1629
90525	AAACGCTCGGTTTATGCTGAATCCGACACCTCGCAGATACTTCT	90574
1629	heProAspGlnValGlnAspLeuValPheAsnLeuHisMetIleLeuSer	1645
90575	TTCCGTAACAAGTACAAAGATTCCTTTAATCTGCATATGATCCCTGCA	90624
1646	AspThrValLysMetLysGlnLysGlnLysAspProGlnMetLeuLias	1662
90625	GATACCTTTAAATGAGAGATACGATACGAAAGACCCGAAATGTCCTTGA	90674
1662	PLeuMetTyrArgIleAlaLysGlyTyrGlnThrSerProGlu...ArgL	1678
90675	CCCTGATGAATCGATTGGCCAGGAGATACCAAAATAACCTGATCTACGAC	90724
1678	euthrTrpLeuGlnAsnMetAlaGlyLysHisSerGluArgSerAsnHis	1694
90725	TGACTTGTTGGAAAAATATGGCTTAATAAACCCGCGAGCAACAAATAC	90774
1695	AlaGlnLysAlaGlnCysLeuValHisSerAlaAlaLeuValAlaGluTyr	1711
90775	ACGGAGCACCGCATGTGTTATGTACATGCTGCTCTTGAATTCGTAATA	90824
1711	rLeuSerMetLeuGluAspArgGlySerTyrLeuProValGlyCysValThrP	1728
90825	TCTTAGCATTTGGAGTACCAAAAAAATTTGCCGTGTGAGCTGTAAAGT	90874
1728	heGlnAsnIleSerSerAsnValLeuGlnGluLysSerAlaValSerAsp	1744
90875	TTCAAGCAATTTCTCCCAACACACTAATGAGATGCGCGTATCGGATGAT	90924
1745	ValValSerProAspGluLysIleCysSerGlyLysTyrPheThrG	1761
90925	GTCTTAAGTCCCGGAGAAATGATATCTCCAGGAAATCAATTCACATCA	90974
1761	uSerGlyLeuValGlyLeuLeuGlnLysAlaIleAspPheSerMet	1778
90975	AACTGGGTTGAAGGCTTCTGGAAGAACCTTCATTTCTTCAAGTTG	91024
1778	IaGlyMetTyrGlnAlaValAsnGluValTyrLysValLeuIleProIle	1794
91025	CTGGCATGTATGAGCAATGAGAGGATGTACAAAAATCTTATATCCGCTA	91074

seq_name: gb_in:AE003590

```
2017 GYLSRYSALGIYThrTValAsnGlnGlyProLeuGluValAlaGlnVa 2033  
||| :::::::::::::::::::::  
91975 GGAGTGATGGACCACTGTTAACCAGAAGCCGATGGAATATGCCAGTCT 92024  
  
2033 IPhelUsSerGlnIleProSerAspProlSylEu..PheArgHisHisA 2049  
92025 GTTCCTTTCCATTATTCGCAGCGAACCAACTGTACCGACAATAACACCAA 92074  
::: :::::  
::: :::::  
2049 snLYSLeuArgLyuCysPhelyAspPheMethLyArgCysGlnuAspLa 2065  
||||| :::::::::::::::::::::  
92075 ACAAACCTTCGGTGTGCTTTCGGGAGCTTTCCAAGCGTTGTGCTCATCT 92124  
||||| :::::::::::::::::::::  
2066 LeuArgLYsaNLYSerLeuileGLyProValGlnLYsgLuTYrgLINar 2082  
||||| :::::::::::::::::::::  
92125 TTGAATAAAAGATCCCAATCTAATACTTTCAGATCAAAAAGATTACCAACG 92174  
|||||  
2082 gGLnuEu 2084  
|||||  
92175 AGAACTG 92181
```

seq_documentation_block:			
LOCUS	AEO03590	301639 bp	DNA linear INV 04-OCT-2000
DEFINITION	Drosophila melanogaster genomic scaffold 14200001386046 section 1 of 16, complete sequence.		
ACCESSION	AE003590 AE002638		
VERSION	AE003590.2		
KEYWORDS	HG.		
SOURCE	fruit fly.		
ORGANISM	<i>Drosophila melanogaster</i>		
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 301639)		
AUTHORS	Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X., Brandon,R.C., Rogers,Y.H., Blazer,R.G., Champe,M., Pfeiffer,B.D., Man,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabot Miklos,G.L., Abril,J.F., Abmayyan,A., An,H.J., Andrews,Pfannkoch,C., Balwin,D., Ballew,R.M., Basu,A., Baxendale,U., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y., Benos,P.V., Bereman,B.P., Bhattacharya,D., Bolintsov,S., Botchan,M.R., Bouck,J., Brockstein,P., Brothier,P., Burris,K.C., Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M., Dodson,K., Dunn,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferrez,C., Ferriera,S., Fleischmann,W., Foster,C., Gabrielian,A.E., Gary,N.S., Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorelli,J.H., Gu,Z., Guan,P., Harris,M., Harris,N.L., Harvey,D., Helman,T.J., Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J., Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpun,G.H., Ke,Z., Kamnitsos,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C., Kravitz,S., Kuip,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A., Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Maizel,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C., Morris,T., Mostreil,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Muszy,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K., Nussernd,D.R., Paclebo,J.T.M., Palazzolo,M., Pitman,G.S., Pan,S., Pollard,J., Puoti,V., Reese,M.G., Reinert,K., Remington,K., Saunders,R.D., Scheeeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I., Simpson,M., Skupski,M.P., Smith,T., Spiter,E., Spradling,A.C., Stapleton,M., Strong,R., Sun,E., Svyrkas,R., Tector,A.C., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassenaar,D.A., Weinstock,G.M., Weissbach,J., Williams,S.M., Woodage,T., Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F., Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H., Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O., Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.		

TITLE The genome sequence of *Drosophila melanogaster*
JOURNAL Science 287 (5461), 2185-2195 (2000)
MEDLINE 20196006
REFERENCE 2 (bases 1 to 301639)
AUTHORS Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
TITLE Direct Submission
JOURNAL Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
COMMENT On Oct 9, 2000 this sequence version replaced gi:7296218.
FEATURES
source
Location/Qualifiers
1..301639
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="2L"
join(2454..3041,3154..3514,3593..4416)
/gene="CG11023"
/product="CT9573"
/db_xref="FLYBASE:FBan0011023"
/db_xref="FLYBASE:FBgn0031208"
<2454..>4416
/map="21A3-21A4"
/db_xref="FLYBASE:FBan0011023"
/db_xref="FLYBASE:FBgn0031208"
join(2605..3041,3154..3514,3593..4201)
/gene="CG11023"
/note="CG11023 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0011023"
/db_xref="FLYBASE:FBgn0031208"
/protein_id="AAF51571.1"
/db_xref="GI:7296280"
/translation="MGERDDPQSSERISIFNPYYTQHOVRNEAPYIPTFDLSDME
ESSORVANAGSPFRLTYSDAVRLSONGFANRSCHSSYTVRRPALUDRSLTSDME
ERMQEOYIYILRTAOSNSYGSHTAPVTDNSKVKYSTNRKLTHTESLHRFASC
VNLNDFRDEQKRAADCEESKRLRLAEQRLEFSNSAIKNIDELRAVERVANI
FQSCIAHRPLFVLGIDPKPNVKVTKLIPTEHHRFNEITDDKSTELIFFENHI
TTEDICTFINGKWLNDVINFPYMSLTERSEKSGVLPATYAINTFEVPLDAGAG
IKRWTRKVDLESKDIIIPVPCNGVHWCMAIHLRNKTIHYDSSKGPKNPVDALEK
YLRESIEFKPKQDSDFVIESYQNIPIRLDSDSCIFSCMAEYITCVPTTFVQS
EMLFRKKMALEIDGEL"
/complement(join(4761..6269,6335..6443,6704..7146,
7211..7853,8445..8550,8608..9799,9858..10636,11978..12137,
13186..134945).8550,8608..9799,9858..10636,11978..12137,
13186..134945)
/gene="1(2)g1"
/product="CT43225"
/db_xref="FLYBASE:FBan0002671"
/db_xref="FLYBASE:FBgn0002121"
complement(join(4761..6269,6335..6443,6704..7146,
7211..7853,8445..8550,8608..9799,9858..10636,11978..12137,
12951..13093,13186..13475))
/gene="1(2)g1"
/product="CT43227"
/db_xref="FLYBASE:FBan0002671"
/db_xref="FLYBASE:FBgn0002121"
complement(<4761..>16291)
/gene="1(2)g1"
/note="CG2671"
/map="21A3-21A4"
/db_xref="FLYBASE:FBan0002671"
/db_xref="FLYBASE:FBgn0002121"
complement(join(4767..6269,6335..6443,6704..7146,
7211..7853,8445..8550,8608..9799,9858..10636,11978..12137,
15990..16291))
/gene="1(2)g1"
/note="Nucleotide sequence of the Celera sequence differs
from the published sequence for this transcript."
/product="CT9059"
/db_xref="FLYBASE:FBan0002671"
/db_xref="FLYBASE:FBgn0002121"
complement(join(6140..6269,6335..6443,6704..7146,
7211..7853,8445..8550,8608..9799,9858..10636,
11978..12061))

/gene="1(2)g1"
/note="1(2)g1 gene product [alt 3]"
/codon_start=1
/db_xref="FLYBASE:FBan0002671"
/db_xref="FLYBASE:FBgn0002121"
/protein_id="AAG22255.1"
/db_xref="GI:10727417"
/translation="MLKPIRGKGOOPSDRHRLOKDLFAYRKTAQHGFPHKPSALAYD
PVILKMAIGTQGALKVFGQPEVEIDGQHTLLNNSASELNOLLEWVIGTGHTLSLA
ANGLIMEPVGATLLPITKLPDGLKIKRYSICSLSDLMIIGEGNTIQDLATF
TIKEPIYHADVLEQVPAYKLNPAIGESIRQLPSPSKLVAAYNRGLCYLMDPSAS
VQRAYIAPGKGSVGLTYVESGEFTWYADDSYATNSIDNEPPSNVYVYGPDC
KSNRLYKGRSRSDIVYVSGMPRSACVHMSDGHKVCIDLPSTKVIDEFTV
FENNRVAEVLVILEELCAVDLPDPCAIKAKYHLISVHSAATCYVLASEVQSV
YESIIRACGEYQIDIDISNIWPTTGGTLPDNLDEESVEDTKLLELIGHEDGSYKFW
DCTGVLLKPIYNEFTSIFGSESDPRDDAADSMEQDEEPPRKSGLDPDSDP
RLAVKRIACFPRTGOLIVGTAGQIVADPFDLPKVKSLKYSIMNLVSDRGGFWKGG
DOLANRSLNLGDEAIPTBERGVNIGVQVLPASITMALIASGLVSGGTANGVLV
EDKRFVPEYFHRCTLNPDLTGAGEQLSRKRSFKSLSESRKLKRGSTRNOSNOV
PITLLEARVERQIEARCADDLGSMWRCLLAKKIYLVNVTISPLMSATNASVSVF
LTLHPAQTAAATVAPSAGNAPPHMRIRISQAKNEIOLKIRAPVVISIDQAGSVY
DOLNAGENGSPPHRYLIIASEBOFKVFSJPLQKLIKYLTLNENGRIRIRHGFESGR
ISPELTQSHGSGSPKTRSHGDGADNIGSLAVSGDVNENFALICLNMMDIV
LSVPELKRQIILAAVREDINGVSLCTNGSEALYHMSSELQRIALATRSVQPPG
VYVPEPLENEESVLENDENKKEITYADDEVVNTYELNPSGISICTRPAEENGRNS
VOOVNGVNIINSNPQANETIISSIDITVDVSVRHLNNTTTTLCISINTEETIGRLSVL
STQIKKASTTYVMSEIPININISNLEDLESKRNTTSTSVVAKSIITINISHEKTNGD
NKIGTPKTAPEESQ"
/complement(join(6140..6269,6335..6443,6704..7146,
7211..7853,8445..8550,8608..9799,9858..10636,
11978..12061))
/note="1(2)g1 gene product [alt 1]; Nucleotide sequence of
the Celera sequence differs from the published sequence
for this transcript"
/codon_start=1
/db_xref="FLYBASE:FBan0002671"
/db_xref="FLYBASE:FBgn0002121"
/protein_id="AAF51570.1"
/db_xref="GI:7296279"
/translation="MLKPIRGKGOOPSDRHRLOKDLFAYRKTAQHGFPHKPSALAYD
PVILKMAIGTQGALKVFGQPEVEYLGQHTLLNNSASELNOLLEWVIGTGHTLSLA
ANGLIMEPVGATLLPITKLPDGLKIKRYSICSLSDLMIIGEGNTIQDLATF
TIKEPIYHADVLEQVPAYKLNPAIGESIRQLPSPSKLVAAYNRGLCYLMDPSAS
VQRAYIAPGKGSVGLTYNPSGEFTWYADDSYATNSIDNEPPSNVYVYGPDC
KSNRLYKGRSRSDIVYVSGMPRSACVHMSDGHKVCIDLPSTKVIDEFTV
FENNRVAEVLVILEELCAVDLPDPCAIKAKYHLISVHSAATCYVLASEVQSV
YESIIRACGEYQIDIDISNIWPTTGGTLPDNLDEESVEDTKLLELIGHEDGSYKFW
DCTGVLLKPIYNEFTSIFGSESDPRDDAADSMEQDEEPPRKSGLDPDSDP
RLAVKRIACFPRTGOLIVGTAGQIVADPFDLPKVKSLKYSIMNLVSDRGGFWKGG
DOLANRSLNLGDEAIPTBERGVNIGVQVLPASITMALIASGLVSGGTANGVLV
EDKRFVPEYFHRCTLNPDLTGAGEQLSRKRSFKSLSESRKLKRGSTRNOSNOV
PITLLEARVERQIEARCADDLGSMWRCLLAKKIYLVNVTISPLMSATNASVSVF
LTLHPAQTAAATVAPSAGNAPPHMRIRISQAKNEIOLKIRAPVVISIDQAGSVY
DOLNAGENGSPPHRYLIIASEBOFKVFSJPLQKLIKYLTLNENGRIRIRHGFESGR
ISPELTQSHGSGSPKTRSHGDGADNIGSLAVSGDVNENFALICLNMMDIV
LSVPELKRQIILAAVREDINGVSLCTNGSEALYHMSSELQRIALATRSVQPPG
VYVPEPLENEESVLENDENKKEITYADDEVVNTYELNPSGISICTRPAEENGRNS
VOOVNGVNIINSNPQANETIISSIDITVDVSVRHLNNTTTTLCISINTEETIGRLSVL
STQIKKASTTYVMSEIPININISNLEDLESKRNTTSTSVVAKSIITINISHEKTNGD
NKIGTPKTAPEESQ"

ANOLIMEPVGATLLPIKILPEDGKLVKSSLOCCSKLDLWIGTEGNTYQDLHTE
 TIREPVIYHDVIVLEOVPAPKINLPALNGAIESIROLNPSKILVANGLCVMDPESAS
 VORAYIAPHGOSGLVNEFSGSEFTWHDGSGYATMSIDNPPSNVNVVPYGPDPG
 KSNIRLYKKKRRSNDVIVFSGMPSRAYGDHNCVSHASGHHKVCIDFJSKVIDEFVT

alignment_scores:

Quality: 4690.00 Length: 2340
 Ratio: 2.886 Gaps: 58
 Percent Similarity: 69.444 Percent Identity: 44.188

alignment_block:

US-09-737-246-2 x AE003590 ..

Align seg 1/1 to: AE003590 from: 1 to: 301639

4 ArgArgAlaPheAlaGlnLysIleSer..... 12
 :::::::::::::::::::::::::::::::
 89795 CAGCGCTGCTTGCGAAGAAAGTTTCCAGTAAGTTTCTACTAAAGCA 89844
 12 12
 89845 ATAGTTTATTTAAACAAGAACGCCCTTAATGCTGTATACATATGTA 89894
 13 ArgThrValAlaAlaGluValArgLys 21
 89895 CATATATCATTTGTAAATCCCTAGGACAGCTCGCTGCGGATGTACGAAAA 89944
 22 GlnIleSerGlyGlnIlyrSerGlySerProGlnIleuLeuLysAsnLeu 38
 :::::::::::::::::::::::
 89945 AATGTTTCTGGTTCATCTATCAAGGCAATG..... GACCCATC 89985
 38 nIleValGlyAsnIleSer.HisHisThrThrValProLeuThr..... 52
 :::::::::::::::
 89986 ATGTGTGGTTCGTTGCGTAAGACACTATGTTTAAATATATACAAATTTT 90035
 53 GluAlaValAspProValAs 59
 90036 CTGTGAAGATTCATTTTTCATATGATCTCCCGACAGAGCCCTCTGA 90085
 59 PleuGlnAspTyrLeuIleThrHisProLeuAlaValAspSerGlyProL 76
 :::::::::::::::
 90086 TTACGAAAGATTTCTATCCAGCATATGAAACATCATCAACCGGATCCT 90135
 76 euArgAspLeuIleGlnPheProProAspAspIleGluValValTyrSer 92
 :::::::::::::::
 90136 TGAACATATATATGACTTCTCAAGGGAGATGTAAGTAAATCAT 90185
 93 ProArgAspCysArgThr.LeuValSerAlaValProGlu..... 105
 :::::::::::::::
 90186 CCCCAGAAAATCAGAACTGTGATCAGTGTCCCAATGAGAAATTTGTA 90235
 106 Glu 106
 90236 TGTATGACACTTAATACACAGCTACATAGCTACAGTCTCTTTAGG 90285
 107 SerGlnMetAspProHisValArgAspCysIleArgSerTyrThrGlnAs 123
 :::::::::::::::
 90286 TGTGATTTGCCCATCAGCTACAGAAATGTGTAACTGATATACGGCCG 90335
 123 pThrAlaIleVal..... 128
 :::::::::::::::
 90336 ATGGAAAGTGTGTAATATGACACAGCGGCACTTCCAGTTCTTGTCTATA 90385
 128 leArgLysTyrHisLysLeuGlyThrGlyPheAsnProAsnThrLeuAsp 144
 :::::::::::::::
 90386 TCCGCGAGGCTATACGACCGGGAAC... ATAGTCCCTCTGCTGCCAG 90432
 145 LysGlnLysGlnuArgLysGlyLeuProLysGlnValPheGlnSerAs 161
 :::::::::::::::
 90433 CAGGAATTCAGATCGACAGAGAT..... TTTTCGTCAAT 90467
 161 pGluAlaProAspGlyAsnSerTyrGlnAspAspGln.. AspAspLeu 177
 :::::::::::::::

90468 TGA..... GAGCGCTTCGCTTATATAAGCGAGAGCTGCACCTCCAAATT 90511
 177 ysArgArgSerMet..... SerIleAspAsp... 185
 :::::::::::::::
 90512 CTCGACAAATCCATTCGAAGTTTGGCTTCAGTAAGCTCTTTACGAGACAT 90561
 186 ...ThrProArgGlySerTyrPheAlaCysSerIlePheAspLeuLysAsn 201
 :::::::::::::::
 90562 TTGACCCCGCGCTGCTCTTGGGCTAGC..... TTTGATCTTCGACGCTTC 90605
 201 rLeuProAspAlaLeuLeuProAsnLeuLeuAspArgThrProAsnGlu 218
 :::::::::::::::
 90606 GGTTAATGACCCACTAATATCCCAATTTGCTTGACAAATGTACCTCCGAGC 90655
 218 IuileAspArgGlnAsnAspAspGlnuArgLysSerAsnArgHisLysGlu 234
 90656 ATATTGATCATCCATGAAAGCAGCGGCGGACGAGATGACAAATGCT 90705
 235 LeuPheAlaLeuHisProSerProAspGluGluGluProIleGluArgLe 251
 :::::::::::::::
 90706 CTTTCTCAGCTGTATCCGAATCTGAACGCGAGGAAACATTTGAACCCCG 90755
 251 uSerValProAspIleProLysGlnHisPheGlyGlnArgLeuLeuValL 268
 :::::::::::::::
 90756 CCGTCTCGACAGAGATTCAGTTGAACATATGGGTCACTGCAATTCAGTTA 90805
 268 yscYleuSerLeuLysPheGluIleGluIleGluProIlePheAlaSer 284
 :::::::::::::::
 90806 ACTGCTACACTTCGCTTGAGCTGAGTGAAGGCCAATTTTTCATTCG 90855
 285 LeuAlaLeuTyrAspValLysGluLysLysIleSerGluAsnPheTy 301
 :::::::::::::::
 90856 ATGGCTATTTATGATGCTAAGGAACGGCAAAATTTCTGAAATTTCTA 90905
 301 rPheAspLeuAsnSerGluGlnMetLysGlyLeuLeuArgProHisValP 318
 :::::::::::::::
 90906 TTTTGACATGAAACGAGATGATCTTAAACGATGTTGCAAGTATATATAC 90955
 318 rProAlaAlaIleThrThrLeuAlaArgSerAlaIlePheSerIleThr 334
 :::::::::::::::
 90956 AGTGGCGGACATATAGTACCAAGGCCACCTCAGCCATATTCGAATTTCA 91005
 335 TyrProSerGlnAspValPheLeuValIleLysLeuGluLysValLeuG 351
 :::::::::::::::
 91006 TATCCGAGTAACGATTTATTTCTGATGTCGTGGAAGAAAGTTTGG.. 91053
 351 nGlnGlyAspIleGlyGluCysAlaGluProTyrMetIlePheLysGluA 368
 :::::::::::::::
 91054 .CAGGCTGACATAAACATTCGCTGAGCGCTATCTC..... 91089
 368 laAspAlaThrLysAsnLysGluLysLeu... GluLysLeuLysSerGln 383
 :::::::::::::::
 91090AAGAAAGCAACAAAGATAATACGAGAGAAAGTAAATTCACAC 91131
 384 AlaAspGlnPheCysGlnArgLeuGlyLysTyrArgMetProPheAlaTr 400
 :::::::::::::::
 91132 GCTGTGAGCTACTGTGAAGCACTAGGCAATATCCGAATGCCGTTTCATG 91181
 400 pThrAlaIleHisLeuMetAsnIleValSer..... 410
 :::::::::::::::
 91182 GACAGGAATCTACATACAAAGTTTAAATGCGGATACATTGAAAGCA 91231
 411 ..SerAlaGlySerLeuGluArgAspSer..... 419
 :::::::::::::::
 91232 AAGAGGCTGGCGCCCGGAGGAGATTCCTTGGAAATGCTTGAGGACG 91281
 420ThrGluValGluIleSerThrLysGlyLysGlySerTr 433
 :::::::::::::::
 91282 AGCTTGGGACAGCGCCAAAGCTCCAAATGACTAGATCGAAATATCTTCAC 91331
 433 pSerGlu..... ArgArg..... AsnSerSerI 441
 :::::::::::::::
 91332 AGTATGCTTTGATCACTAAGCGGAAAGCTAAGCATATGAGCGGACAC 91381

1596	ValGlyThrSerGlnAsnProPheAsnGluGluPheLeuArgArgSerLeu	1612
94764	GTGGAAACAAGTTCGGCTTTAGTGAACACCTCTTACGGCGTCGGCTTAA	94813
1612	SThrIleLeuThrTyrArgIleGluGlnAspLeuGluLeuArgGluThrThp	1629
94814	AACCGTCTCGTGGTTATCTCTGAATCCGACCTCCGACCTGCAGAGATCTT	94865
1629	heProAspGlnValGlnAspLeuValPheAsnLeuHisMetIleLeuSer	1645
94864	TTCCGTAACAAGAACAAAGATTCTCTTTTAATGCAATGATCCGTGA	94914
1646	AspThrValIysMetIysGluHisGlnGlnAspProGluMetLeuLeu	1662
94914	GATACGTCTTAAATGAAAGAGACTACAGAGAACCCAGAAATGTCCTTGA	94966
1662	PleuMetTyrArgIleAlaIysGlyTyrGlnThrSerProGlu...ArgL	1678
94964	CCCTCATAACTCGATTTGCCAAAGGAAATCCAAATAATCCGTGATCTAGCAG	95013
1678	euthrThrPleuGlnAsnMetArgIlyLysHisSerGluThrSerAsnHis	1694
95014	TGACTTGTTGGAAATATGGCTTAAAAAACCCGACGAGCAAAATTCAC	95063
1695	AlaGluAlaIleAlaGlnCysLeuValHisSerAlaAlaLeuValAlaGlu	1711
95064	ACGGAACACCCATGTTGTGTGACATGCTGCTTTGTTAGTTCTGATA	95113
1711	rIleuSerMetLeuGlnAspArgIysTyrIleuProValIglyCysValThp	1728
95114	TCTTAGCATTTGGAGTCAACAAAACATTTGCTGTTGGAGCTGTAAGTT	95163
1728	heGlnAsnIleSerSerAsnValLeuGluGluSerAlaValSerAspAsp	1744
95164	TTCAACAATTTCTCCAAACACACTAATGAGTGGCGCGATACGATGAT	95213
1745	ValValSerProAspGluGluGlyIleCysSerGlyLysTyrPheThrGlu	1761
95214	GTCGTAAAGTCCGGCGAAGATGGTATCTGCTGGAAGAAATCATTTCACTGA	95263
1761	userGlyLeuValGlyLeuLeuGluGlnAlaAlaIleSerPheSerMetA	1778
95264	AACGGGTTAAAGCCCTGCTGTGAAAGAACCTCCATTTCTTCAAGTGG	95313
1778	IaGlyMetTyrGluAlaValAlaGlnGluValTyrLysValLeuIleProIle	1794
95314	CTGGCATGTTGTGAAGCAATGACGAAGTGTACAAATTTCTAATACCATA	95363
1795	HisGluAlaAsnArgAspAlaLysLysLeuSerThrIleHisGlyLysIle	1811
95364	TGGAGGCTTACAGACGATTTTCAAAAGCTAAGCAAAATTCAAGCAAAAT	95413
1811	uGlnGluAlaPheSer.....	1816
95414	GCAGGAGGCAATTAAATCAATATCCCACTACAGTAAGTAATATTGTGA	95463
1817LysIleValHisGlnSerThrGlyTyrPheGluArgIle	1828
95464	AATTTTACCAACGGAATAATATATACATATTATTAACACAGGTAAAGAGT	95513
1828	tpheGlyThrTyrPheArgValGlyPheTyrGlyThrLysPheGlyAspL	1845
95514	TTTTGGACACTACTTCGCTTGCTGCTTATGCGGAAAAATTTGGGACT	95563
1845	euaSpGluGlnGluPheValTyrLysGluProAlaIleThrLysLeuAla	1861
95564	TGATATCAGCAGGAATTCATTATTAAGAGCCAAACATTGACGAAGTGGCC	95613
1862	GluIleSerHisArgLeu...Glu.....	1868
95614	GAATATTTTAGTGCCTTACAGGATATATATTCGAATTTGAAAAATATAGA	95663

1869 GlypHeTyrglyGluArgPhe 1875
95664 CTAATCAATTTTGTTCACATACGTTAGAACTTTTACCTGACGACATTC 95713
1876 GlyGluAspValValGluValIleLysAspSerAsnProValAspLysCys 1892
||| ||| |||||:::|||||:::|||||
95714 GCACCGGACCTCTGTGCATATCATTTAAAGATTTCCATATCCCTGTGATTTAA 95763
1892 sLysLeuAspProAsnLysAlaTyrlIleGlnIleThrTyValGluProT 1909
:::|||||:::|||||:::|||||:::|||||
95764 TAGCTTGATCCCGATTAAGGCTTCATTTCAATTAATTACTATGTGAACCT 95813
1909 yRheAspRhlThrTyGluMetLysAspArgIleThrTyRheAspLysAsn 1925
95814 ACTTTGAACATATGAATGCGTCATGTCGAGACATCTTTGAGCCGAAT 95863
1926 TyrAsnLeu..... 1928
:::|||||:::
95864 TTCAATTAATTAAGTATGATATGATTAACACAGTAAATTTAAATCCAAATTTA 95913
1929Arg.ArgPheMetTyRys 1934
||| |||||:::|||||:::
95914 AAATTAATATTTTACATTTTGGTTTATGCAGACAGCTTTATATATATGCC 95963
1935 ThrProHehrhLeuAspGlyArgAlaHisGlyGluLeuHisGluGlnPh 1951
|||||:::|||||:::|||||:::|||||
95964 ACGGCTTTTACTTAAAAATGGAAAGCCACATGCGGAACCTTAATAGCGCTG 96013
1951 eLysArgLysThrIleLeuThrHisSerHisAlaPheProTyrlIleLysT 1968
|||||:::|||||:::|||||:::|||||
96014 TAAACGAAAAACGATTTTGCACCGGCGGATTCATCTTCCATAGCTATAAAA 96063
1968 hArgValAsnValThrHisLysGluGluIleIleLeuThrProIleGlu 1984
|||||:::|||||:::|||||:::|||||
96064 CTCGATTAATGTGTAATAGCCGACAGCAAAATAGTCGTGAACCGATTTGA 96113
1985 ValAlaIleGlu..... 1988
|||||:::|||||
96114 GTGGCAATTTGAACTAAGTTATGTGTAAATCCCAAAATTAATTAATTAAT 96163
1989AspMetGlnLysLysThrGlnGluLeuAlaPheA 2000
|||||:::|||||:::|||||
96164 AAACGTTATTCCTAGGATATTTCAAAAAAAAACATTTGGAATTTGGCAAGCTG 96213
2000 lathrHisGlnAspProAlaAspProLysMetLeuGlnMetValIleGln 2016
|||||:::|||||:::|||||:::|||||
96214 CCAGCAATCAAGAAGCCAGACGCCCAAAATATTTGCAATAGTACGACAG 96263
2017 GlySerValGlyThrThrValAsnGlnGlyProLeuGluValAlaGlnVal 2033
||| ::|||:::|||||:::|||||:::|||||
96264 GGGATGATTTGGAACCACTGTTAAACCAAGGCCCATTAATGGCGAGTGT 96313
2033 lPheLeuSerGluIleProSerAspProLysLeu...PheArgHisAla 2049
|||||:::|||||:::|||||
96314 GTTCCCTTTCCAAATTTATCCGACGGAACAACTGTAACGACAAACACCAAA 96363
2049 snLysLeuArgLeuCysPheLysAspPheThrLysArgCysGluAspAla 2065
|||||:::|||||:::|||||:::|||||
96364 ACAACATCGCTGTGCTTCCGCGAGTTTCAAGCGTGTGTGCTANTGCT 96413
2066 LeuArgLysAsnLysSerLeuIleGlyProValGlnLysGluTyrlnAr 2082
|||||:::|||||:::|||||
96414 TTGAAAAAAGATCGCAATTAATACCTTTCAGATCAAAAAGATTAACCAAG 96463
2082 gGluLeu 2084
|||||
96464 AGAAGCTG 96470
seq_name: gb_pr:AK055905

to Rat trng gene product.
 AK055905
 AK055905.1 GI:16550749
 Oligo capping: f1s (full insert sequence).
 Homo sapiens normal mesangial cells (NHMC56046-2) cDNA to mRNA,
 clone_110: MESANI000101.
 ORGANISM
 Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 1 (sites)
 Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
 Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R.,
 Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
 Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
 Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
 Magatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
 Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
 Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 3327)
 Isogai,T., Otsuki,T. and Sugiyama,T.
 Direct Submission
 Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
 Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 252-0812, Japan
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan: cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: RAB and
 HRI.
 FEATURES
 source
 Location/Qualifiers
 1..3327
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MESANI000101"
 /cell_type="normal mesangial cells (NHMC56046-2)"
 /clone_lib="MESANI"
 /note="cloning vector: pME18SFL3-primary culture, normal
 mesangial cells"
 54..2279
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB71042.1"
 /db_xref="GI:16550750"
 /translation="MFELSVFROOHYLAGIVLTAVLIDPDAGSLFGLHKYINW
 HNLSHSDPRVSDPOIKAVAMLYLIGIIMETVPOLYDFETINORRPICLAT
 DDYSESGMSISOTVAMVAGTVPOLRPGSFLLTSGROHTTFEASRSLLICL
 LMVKNDEVTYLOKMPDLSVLOINRLIDILYLCVSGFEKKGKGRPMNSLTPKSK
 DMRAKLEPAITIGSTGAREMYRBSRGOLRPSGARSAGSDNLEMRDMDHMRNTEK
 LDKRAETHEALIDGNLATRANILITIDLETIVQYVSIESKSIIGCVILYLSHK
 ACNOSAVYLQHCFAFQALVRSPELLLEETECQACJLNLPHCSSIGTIRSHS
 ASLYLMRNOEIGNNFARVQMVTMSLSVGSQDNNEFLRSLKTLITVAEHL
 ELRETFEPDOVDVLEHMLITSDYVAKKEHOEPENMLIDMTYIAGYQSPDLRLT
 WLONNKGHSERSNHAEAOCLIVSAALVAEYLSMLEBKRLTPGCVTFQVRSNVE
 ESAYSDDVSPDEGICSGKRYTESGLVLEOAAASNMAQWEANEVYKXVILPIH
 EANDRAKLLSTHKLQDPAFSKTYHODSKRRFGYFRPGFTGTFGLDDDEEPYKRP
 AITKLAEISHLRLEGFYGERFGEDEVVEYIKDSNPVCKLDENHIFRLPMSHRTLTH
 R"

Align seg 1/1 to: AK055905 from: 1 to: 3327
 1156 SerGlnSerSerGlyPheSerThrAsnValGlnAspGlnLysIleAlaAs 1172
 |||||||
 3 TCTGAGATTCGTGATTTCTTCTAGCAATGTACAGACCAAAAGATTGCCAA 52
 1172 nMetPheGlnLeuSerValProPheArgGlnGlnHisTyrLeuAlaGlyL 1189
 |||||||
 53 TATGTTGAATATTCGTCGCTTCGCCCAACAGACATTAATTGGCAGAGAC 102
 1189 euValIleuThrGlnLeuAlaValIleLeuAsProAspAlaGluGlyLeu 1205
 |||||||
 103 TTGTGTTAACGAGAGCTGCGTGTCTTTTAAACCTGATGCTGAAGGACTG 152
 1206 PheGlyLeuHISLysLysValIleAsnMetValHisAsnLeuSerSe 1222
 |||||||
 153 TTGGATTGCATTAAGAAAGCATCATATATGATGACACATTTACTCTTCAG 202
 1222 rHisAspSerAspProArgTyrSerAspProGlnIleLysAlaArgValA 1239
 |||||||
 203 TCAGAGACTCAGACCCGGGCTACTGTGACCCCTCAGATAAAGGCTGAGTGC 252
 1239 IamettLeuTyrLeuProLeuIleGlyIleIleMetGluThrValProGln 1255
 |||||||
 253 CCAGTGTGATCTACCTCTGATGTTGATTCATGAGAACTGATACCTCAG 302
 1256 LeuTyrAspPheThrGlnThrHisAsnGlnArgGlyLysProIleCysII 1272
 |||||||
 1272 eAlaThrAspAspTyrGlnSerGlnSerGlySerMetIleSerGlnThrV 1289
 |||||||
 353 AGCCACTGATGATTATGAAAGTGAGAGGGAAGTATGATAAGCAGACCG 402
 1289 aAlaIleMetAlaIleAlaGlyThrSerValProGlnLeuThrArgProGly 1305
 |||||||
 403 TTGCCATGGCAATCGCAGGACATCGCTCCCTCAACAAAGGCTGCGC 452
 1306 SerPheLeuLeuThrSerThrSerGlyArgGlnHisThrThrPheSerAl 1322
 |||||||
 453 AGTTTCCTCTCAGCTCAACGAGTGCGGCAACACACTACTTTCAGC 502
 1322 agLusSerSerArgSerLeuLeuIleCysLeuLeuTyrValLeuLysAsnA 1339
 |||||||
 503 AGAATCAAGTCGAAACCTTTGATCTGCTACTTGGGTTGCCAAATG 552
 1339 IAspGluThrValLeuGlnLysThrPheThrAspLeuSerValLeuGln 1355
 |||||||
 553 CAGATGAACACAGTTCTACAGAAAGTGTTACAGATCTCTCAGCTTTCAG 602
 1356 LeuAsnArgLeuLeuAspLeuLeuTyrLeuCysValSerCysPheGluTy 1372
 |||||||
 603 CTAACCCGCTATTAGATCTGCTTATCTCTGTGTCGTCTTTCAGTA 652
 1372 rLysGlyLysLysValPheGluArgMetAsnSerLeuThrPheLysLys 1389
 |||||||
 653 TAAAGGAAAAAAGTGTGAACGAATGAATAGCTTGACCTTTAAGAAAT 702
 1389 eLysAspMetArgAlaLysLeuGlnGluAlaIleLeuGlySerTleGly 1405
 |||||||
 703 CAAAGACACTGAGACCAAGCTTGAAAGACATTTCTTGGAGGCTTAGGT 752
 1406 AlaArgGlnGluMetValArgArgSerArgGlnLeuGlnArgSerPr 1422
 |||||||
 753 GCCAGGCAAGAAATGTAGTACGGGAGAGCCGAGACACGTCGAGAGAGCC 802
 1422 oSerGlySerAlaPheGlySerGlnGlnAsnLeuArgTyrParGlyAspm 1439
 |||||||
 803 ATCTGGAAGTGCCTTTGGAAAGCAAGAAAAATTTGAGGCGGAGAAAGTA 852
 1439 eTThrHisTyrParGlnAsnThrGlnLysLeuAspLysSerArgAlaGlu 1455
 |||||||

853 TGACTCACTGGCGCTCAAAACACAGAGAGCTTGCAATCAAGACGAGAG 902
 1456 ILEGLIHISGLIUALALEUILEASPGLYASNLEUVALATHRGUUALASOLE 1472
 903 ATTGAACACGAGCACTGATGTGATGAAACCTGGCTACAGAAAGCAACCT 952
 1472 UUILEILEUASPTHRLLEUGLUILEVALVALGINTHRVALSERVALTHRG 1489
 953 AATCATTTTAGATACATTTAGAGATTTGTTGACAGACGTTTCGTAAACGG 1002
 1489 IUSERTYSGIUSERIILEUGLYGLYVALLEULYSVALLEUHISSER 1505
 1003 AATCCAAAGAGACATTTCTGGGAGCTCTAAAGTGTCTACTACACACC 1052
 1506 METALCYASNGLINSERALAVALTYRLEUGLHISCYSPHEALATHRG 1522
 1053 ATGGCTGTAAACCAAGTGCAGTTTATCTACAACTGTTTGTCTACACA 1102
 1522 NATGALALEUVALSERLYSPHEPROGLULEUENPHEGLUGLUTHRG 1539
 1103 GAGAGCCTTGCTTCAAGCTTCTCTGAACTCTTATTTGAAGAAGACAG 1152
 1539 IUGINCYSALAASPLEUCYSLEUARGLEUENARGHISCYSSERSESR 1555
 1153 AGCAGTGTCTGATTTATGCTCAGGCTTCTCCGACACTGTAGCAGTACG 1202
 1556 ILEGLYTHRILEARGSERHISPROSERALASERLYRLEULEUMETAR 1572
 1203 ATCGGTACATACGGTCACACGCCAGTGCTCCCTTTACTCTATATGAG 1252
 1572 GGLINASNPHGLIILEGLYASNASNPHLEALARGVALLYSMETGLNALP 1589
 1253 GCMAAACITTTGAGATGGGAAATACCTTGCCAGGGTTAAATGACAGTAA 1302
 1589 ROMETSERSESRSESRLEUVALGLYTHRSERGLINASNPHASNGLUGLU 1605
 1303 CATGTGCATATACCTCTTGGGGACATCTCAGAAATTTTATATAAATA 1352
 1606 PHELEUARGARGSERLEUYSRTHRIILEUUTHRYALAGLUUASPLE 1622
 1353 TTCTTAAGACGTTCTCTAAAGACTATATGACATATGCTGAAGAAGACT 1402
 1622 UGLULEUENARGLUUTHRTHRPHERPASPGLINVALGINASPLEUVALPHEA 1639
 1403 GGAATTTGAGGAACAAACATTTCTCGATCAGATCCAGGATCTGGTTTCA 1452
 1639 SNLEUHISETIILEUSERASPTHRVALLYSMETLYSGLUHISGLINGLU 1655
 1453 ATCTCATATGATTTCTTCTGATCTGAAATGAAAGAACACCGAGAG 1502
 1656 ASPPROGLUMETLEUILEASPLEUMETLYRARGILEALALYSGLIYTRGI 1672
 1503 GATCTCTGAATGTGATTTGATCTATGTAGAAATTTGCCAAGGGTTACCA 1552
 1672 NTHRSERPROGLU... ARGLEUTHRTPLEUGLINASNMETLAAGLYLYSH 1688
 1553 GACCTCTCAGATCTGCAGACTGACCTGGTTGCAGAAACATGGCAGGCAAGC 1602
 1688 ISSERGLIARGSERASNHISALAGLUUALAGLINCYSLEUVALHISER 1704
 1603 ACTCGAAGCAAGCAATCATGCTGAAGCTGCACAGTGTCTAGTCCACATCA 1652
 1705 ALAALALEUVALALAGLUYRLEUSERMETLEUGLUNSPARGLYTYRLE 1721
 1653 GCAGACATTTGTTGCTGATATTTTGAGCATGCTGGAGGACCGGAATATCT 1702
 1721 UPROVALGLYCYVALTHRPHGLINASNIISSERSERASNVALLEUGLUG 1738
 1703 TCCGTGGGAGGTGTAAACATTTCAAAATATTTCACTAATGTTTAAAG 1752
 1738 IUSERALIVALSERASPPVALVALSERPROASPGLUUGLYLILCYLS 1754
 1753 AATCTGCGGCTCTCAGATGATGTGTATCTCCAGATGAAGAAGGATATCTGC 1802

1755 SERGLYLYSTRYPHETHRGUSERGILYUVALJILYLEULEUGLUGLINAL 1771
 1803 TCTGGAAAATACTTACTAGATCAGGACTTGTTGGATTACTGGAAACAAC 1852
 1771 AALALASERPHESERMETALAGLYMETTYRGLUALAVALSNDLUVALT 1788
 1853 AGCTCCTTCTCTCTYATGGCTGCATGTATGAAGCAGTTTAATGAATTT 1902
 1788 YRLYSVALLEUILEPROLIEHISGLIUALAASNARGASPALALYSLSLEU 1804
 1903 ACAAGTACTTATTCTATTCATGAACCTAATCGGATGCAAGAAACTA 1952
 1805 SERTHRIIEHISGLYLSLEUGLUALALAPHERLYSTILEVALHISGLI 1821
 1953 TCCACAAATTCATGGTAAACCTTCAAGAAGCANTCAGCAAAATTTGTCTCA 2002
 1821 NSETRHTRGLYTRPGIUAARGMETPHEGLYTHRTRYRPHARGVALGLYPHER 1838
 2003 GGAT.....GGTAAAGCGAGATGTTGGCACCTATTTTCGTGTGGTTTTT 2046
 1838 YRGLYTHRLYSPHGLYASPLEUASPGLUUGLUPHEVALTYRLYSGLU 1854
 2047 ATGGAACCAAGTTTCGGGATTTGGATGAACAAGAAATTTGTTACAAGAG 2096
 1855 PROLAILETHRLYLSLEUALAGLUILESERHISARGLEUGLUGLYPHERY 1871
 2097 CCTCAATTAACCAACTTGCAAGATATCTCACGATTTGAGGGAATTTTA 2146
 1871 RGLYLUARGPHEGLYGLIUAASPYVALJILUVALIILELYASPSERASNP 1888
 2147 CGGAGAAGAATTTGGAGAGATGTGTTGAAGTAAATCAAGACTCTAATC 2196
 1888 ROVALASPLYCYLSYLSLEUASPPROASNLYSALARYLILEGINIETHR 1904
 2197 CTGTAGACAAGTGTAAATTAGATCCCTAACAC..GCATATATTCAGATTAC 2245
 1905 TYRVALGLUPROTYRPHASPTHRYRGLUMETLYSASPARGLIETHRY 1921
 2246 TATGTGAGCCATPACTTTGACACATATGATGATAAGCAAGAAATCACTTA 2295
 1921 RPHEASPLYASNTYRASNLEUARGARGPHEMETRYCYSTHRPROPHET 1938
 2296 TTTTCGACAAAATTTACAACTCTTCGTCATTCATGTACTGTAACCCCTTAA 2345
 1938 HRLEUASPGLYARGALAHISGLYGLULEUHISETLUGLINPHELYSARGLYS 1954
 2346 CTTTAAATGGCCCGCCCATGGGGAACCTTCATGAACAATTCAAAAGGAG 2395
 1955 THRILEUUTHRTHRSERHISALAPHEPROTYRILELYSRHARGVALAS 1971
 2396 ACCATTTGTGACTACGTCTCATGCTCTTCTTATATTAANAACAAGGGTCAA 2445
 1971 NVALTHRHISLYSGLUUUILEIILEUUTHRPROILEGLIUALALILEG 1988
 2446 TGTACACTCATAAAGAAGAGATCATCTTAACACCAATTTGAAGTTGCTAATG 2495
 1988 IUSPMETGLINLYSTRHGLINGLULEUALAPHEALATHRHISGLINASP 2004
 2496 AGGACATGCAAGAAAAAGACACAGAGTTGGCATTTCAACACATCAGGAT 2545
 2005 PROLAASPPROLYSMETLEUGLIMETVALLEUGLINGSERVALGLIYTH 2021
 2546 CCGGACAGCCCCAAAATGCTTCAAGATGTRACTCAAGGATCTGTAGGCAC 2595
 2021 TRHVALASNGLINGLYPROLEUGLUVALALAGLINALPHEUUSERGLUT 2038
 2596 CACAGTGAATCAGGGCCCTTTGGAAGTTGCCACGGTTTTCGTCTGAAA 2645
 2038 IEPROSERASPPROLYSLEUPHEARGHISHSANLYSLEUARGLYEUCYS 2054
 2646 TACCTAGTGACCCAAAGCTCTTCAGACATCATTAATAAATGCGAGCTTGC 2695

2055 PheLysAspPheThrLysArgCysGluAspAlaLeuArgLysAsnLysSe 2071
|||||
2696 TTTAAGATTTTACTAAAGGTGTGACATGCTTAAAGAAAAATAAGAG 2745
|||||
2071 rLeuIleGlyProValGlnLysGluTyrGlnArgGluLeu..GlyLysLe 2087
|||||
2746 CTTAAATGGCGCGGATCAAAAGAGATCAAAAGGAACTGGAGAGAAACT 2795
|||||
2087 userSerPro 2090
|||||
2796 ATCATCGCCT 2805
|||||
seq_name: gb_pat: AX172978
seq_documentation_block:
LOCUS AX172978 3253 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 105 from Patent WO0142297.
ACCESSION AX172978
VERSION AX172978.1 GI:14597957
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 3253)
AUTHORS Lu, P., Garman, J.D. and Candia, A.F.
TITLE Clasp-3 transmembrane protein
JOURNAL Patent: WO 0142297-A 105 14-JUN-2001;
Arbor Vita Corporation (US)
FEATURES
Source 1..3253
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="polynucleotide fragment"
BASE COUNT 1009 a 633 c 705 g 906 t
ORIGIN
alignment_scores:
Quality: 4626.50 Length: 904
Ratio: 5.123 Gaps: 1
Percent Similarity: 99.889 Percent Identity: 99.558
alignment_block:
US-09-737-246-2 x AX172978 ..
Align seg 1/1 to: AX172978 from: 1 to: 3253
1188 GlyLeuValLeuThrGluLeuAlaValIleLeuAspProAspAlaGluG 1204
|||||
4 GGACTTGTTGTTTTCAGAGCTGCTGTCATTTTGAACCCCTGATGCTGAAG 53
1204 yLeuPheGlyLeuHisLysLysValIleAsnMetValHisAsnLeuLeu 1221
|||||
54 ACCTGTTGATTTGCATAGAAAGTCATCAATATGTGACACAAATTTACTCT 103
1221 eSerHisAspSerAspProArgTyrSerAspProGlnIleLysAlaArg 1237
|||||
104 CCAGTCACGACTCAGACCCGCGTACTCTGACCCACGATTAAGCTCGA 153
1238 ValAlaMetLeuTyrLeuProLeuIleGlyIleIleMetGluThrValPr 1254
|||||
154 GTGGCCATGTTGATCTTACTCTGATGTGATATCATGTGAACCTGTACC 203
1254 oGlnLeuTyrAspPheThrGluThrHisAsnGlnArgGlyArgProIleC 1271
|||||
204 TCAGCTGTTGATTTTACGAAGCTCACAATCAACGAGAGAACCAATTT 253
1271 ySLeIleAlaThrAspAspTyrGluSerGluSerGlySerMetIleSerGln 1287
|||||
254 GTATACCACTGATGATTAAGAAGTGAGAGCGGAAGTATGATTAAGCCAG 303
1288 ThrValAlaMetAlaIleAlaGlyThrSerValProGlnLeuThrArgPr 1304
|||||

304 ACCGTTGCCATGGCAATCGCAGGACATCGGTCCTCAACATAAGAGCC 353
1304 oGlySerPheLeuLeuThrSerThrSerGlyArgGlnHisThrThrPheS 1321
|||||
354 TGGCAGTTTCTCTCTCAACGTCACAGAGTGGCAGGCAACACACTACTTTT 403
1321 eAlaGlySerSerArgSerLeuLeuIleCysLeuLeuThrValLeuLys 1337
|||||
404 CAGCAGAAATCAAGTCGAGAGCTTTTGATCTGTCTACTTTGGGTTCTCAA 453
1338 AsnAlaAspGluThrValIleGlnIlySTPrPheThrAspLeuSerValIle 1354
|||||
454 AATGCAGATGAAACAGTTTCTACAGAAAGTGTTCAGATCTCTCAATCTT 503
1354 uGlnLeuAsnArgLeuLeuAspLeuLeuTyrLeuCysValSerCysPheG 1371
|||||
504 GCACATTAAACCGGCTATTAGATCTGCTTTATCTCTGTGTCTTGGCTTGG 553
1371 LutyTyrGlyLysLysValPheGluArgMetAsnSerLeuThrPheLys 1387
554 AGTATTAAGGGAAGAAAGTGTTCAGAGAAATGAATGACTTGACCTTTAAG 603
1388 LysSerLysAspMetArgAlaLysLeuGluGluAlaIleLeuGlySerI 1404
604 AAATCAAAAGACATGAGAGCAAGCTTGAAGAGCTATCTTGGAGAGAT 653
1404 eGlyAlaArgGlnGluMetValArgArgSerArgGlyGlnLeuGluArgS 1421
654 AGGTGCCAGGCCAAGAAATGCTACGGCCAAACCCAGGACAGCTGAGAGAA 703
1421 eProSerGlySerAlaPheGlySerGlnGluAsnLeuArgTyrParGlys 1437
704 GCCCATCTGGAAGAGCTTGTGAAGATCAAGAAATTTGAAGTGAAGAGAA 753
1438 AspMetThrHisTrpArgGlnAsnThrGluLysLeuAspLysSerArgAl 1454
754 GATATGACTCAGCTGGCGTCAAAACACAGAGAGACCTTGACAAATCAAGAGC 803
1454 aGluIleGluHisGluAlaLeuIleAspGlyAsnLeuAlaThrGluAla 1471
804 AGAATATTGAACACGACACACTGATTTGATGGAACCTGTGCTACGAAGCA 853
1471 snLeuIleIleLeuAspThrLeuGluIleValValGlnThrValSerVal 1487
854 ACCTAATCATTTTATAGATACATTAGAGATTGTGTGACAGCCGTTCTGTA 903
1488 ThrGluSerLysGluSerIleLeuGlyGlyValLeuLysValLeuLeuH 1504
904 ACGGAATCCAAAGAGACATTTCTGTGAGAGTGTAAAGTGTACTACA 953
1504 sSerMetAlaCysAsnGlnSerAlaValTyrLeuGlnHisCysPheAla 1521
954 CAGCATGGCGCTGTAAACCAAGTGCAGTTTATCTTAACACACTGTTTCTTA 1003
1521 hrGlnArgAlaLeuValSerLysPheProGluLeuLeuPheGluGlu 1537
1004 CACAGAGAGCTTGTGTTCAAAAGTTCTGAACTTATTTGAAGAGAG 1053
1538 ThrGluGlnCysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSe 1554
1054 ACAGAGCAGTGCTGATTATTAAGCTCAGAGCTTCTCCGAGACGTGTACAG 1103
1554 rSerIleGlyThrIleArgSerHisProSerAlaSerLeuTyrLeuLeu 1571
1104 TAGCATGGTTCATATAGGTACACCCAGAGCTCCTTACTACTATA 1153
1571 eTArgGlnAsnPheGluIleGlyAsnAsnPheAlaArgValLysMetGln 1587
1154 TGAGGCAAAACTTTGAGATTGGGAATTAACCTTGCAGAGGTTAAATGAC 1203
1588 ValProMetSerLeuSerSerLeuValGlyThrSerGlnAsnPheAsnG 1604
1204 GTACCAATGTCTATCTCTCTGTGGGACATCTCAGAAATTTTAAATGA 1253

1604 uGIupheueArGArSerLeuLysThrIleuThrTrAlaGluGUA 1621
 1254 AGAATTCTTAAGACGTTCTCTAAGACTATATTGACATATGCTGAAGAG 1303
 1621 sPLeuGIuLeuArGluThrThrPheProAspGlnValGlnAspLeuVal 1637
 1304 ATCTGGAAATTGAGGGAACACATTTCTCGATGACGTCAGGATCTGGTT 1353
 1638 PheAsnLeuHISMetIleuSerAspThrValLysMetLysGluHISG 1654
 1354 TTCATCTTCATATGATTTCTTCTGATACGTGAAATGAAAGAACACCA 1403
 1654 nGIuAspProGIuMetLeuIleAspLeuMetLysArgIleAlaLysGlyT 1671
 1404 GGAGATCTCTGAATGTTGATGATCTATGATGACGAATTCGCAAGGTT 1453
 1671 yGIuThrSerProGlu...ArgLeuThrTrPLeuGlnAsnMetAlaGly 1686
 1454 ACCAGACCTCTCCAGATCTGCGATTGACCTGGTTCAGAACATGGCAGGC 1503
 1687 LysHISerGIuArgSerAsnHISAlaGluAlaGlnCysLeuValHIS 1703
 1504 AAGCACTCGAAGCAGACGATCATGCTGAGCTGCACAGTGTCTAGTCCA 1553
 1703 sSerAlaAlaLeuValAlaGluTrPLeuSerMetLeuGluAspArgLysT 1720
 1554 CTCAGACAGACTGTTGGTGATATTTCAGCATGCTGAGAGACCGGAAAT 1603
 1720 yTrLeuProValGlyCysValThrPheGlnAsnLISerSerAsnValLeu 1736
 1604 ACCTTCCTGTGGATGTGTAAACATTTTCAGAAATATTTCATCTAATGTTT 1653
 1737 GIuGluSerAlaValSerAspAspValValSerProAspGIuGluGlyI 1753
 1654 GAAGATCTGCGGCTCTCAGATGATGGTATCTCCAGATGAAGAGGAT 1703
 1753 eCysSerGIuLysTrPheThrGIuSerGIuLeuValGlyLeuLeuGluG 1770
 1704 CTGCTCTGAAATACTTACTGACTCAGGACTGTGGATTTACTGGAAC 1753
 1770 LnaAlaAlaLISerPheSerMetAlaGlyMetTrYGluAlaValAsnGlu 1786
 1754 AAGCAGCTGCTTCTCTCTATGCTGCATGTATGAAGCAGTTAAATCA 1803
 1787 ValTrYLysValIleuIleProIleHISGluAlaAsnArgAspAlaLys 1803
 1804 GTTTCACAAAGTACTTATCTTATTCATGAACTAATCGGATTCGAAGAA 1853
 1803 sLeuSerThrIleHISGlyLysLeuGlnGluAlaPheSerLysIleValH 1820
 1854 ACTATCCACAATTCATGCTAAACTTCAAGAGCAATTCAGCAAAATTTG 1903
 1820 ISGlnSerThrGIuTrpGIuArgMetPheGIuThrTrYPheArgValGly 1836
 1904 ATCAAGAGTACGGCTGGGAGCGGATTTGGCCACTAATTTTCGTTGGT 1953
 1837 PheTrYGlyThrLysPheGIuAspLeuAspGIuGlnGluPheValTrYLys 1853
 1954 TTTTATGAACCAAGTTGCGGATTTGGATGATACACAAGATTTGTTTCAA 2003
 1853 sGIuProAlaIleThrLysLeuAlaGluIleSerHISArgLeuGluGlyP 1870
 2004 GGAGCTCTCAATTAACCAAACTTGCAGAGATATCTCACAGATGGAGGAT 2053
 1870 heTrYGlyGIuArgPheGIuLysAspValValGluValIleLysAspSer 1886
 2054 TTTACGGAGAAAGATTGGAGAGATGGGTGAAGTAAAGACTCTT 2103
 1887 AsnProValAspLysCysLysLeuAspProAsnLysAlaTrYLISGlnI 1903
 2104 AATCTGTAGACAAAGTGAATTAGATCTTAAACAAGCATATATTACGAT 2153

1903 eThrTrValGluProTrYPheAspThrTrYGluMetLysAspArgIleT 1920
 2154 TACCTATGTGGAGCCATACCTTTTGACACATATGAGATGACAGACAGAT 2203
 1920 hrTrYPheAspLysAsnTrAsnLeuArgArgPheMetTrCysTrhPro 1936
 2204 CCTATTTCGACAAAATTTACAAATCTTCGATTCGATTCATGCTGACACCC 2253
 1937 PheThrLeuAspGIuArgAlaHISGlyGluLeuHISGluGlnPheLysAr 1953
 2254 TTTACTTTAAGTGGCGGTGCCCATGGGAACTTCATGAACAATTCAAAAG 2303
 1953 glySThrIleLeuThrTrSerHISAlaPheProTrYLISLysThrArgY 1970
 2304 GAACACCATTTCTGACTGCTGCTCATGCTTCTTATATTAAAAACAAGG 2353
 1970 aLAsnValThrHISLysGluGluIleIleLeuThrProIleGIuValAla 1986
 2354 TCATGTCACTCATAAAGAGAGATCATCTTAAACCAATTTGAAGTTGCT 2403
 1987 IleGluAspMetGlnLysLysThrGlnGluLeuAlaPheAlaThrHISG 2003
 2404 ATTGAGGACATGCGAAGAAAGACACAGAGATGGCATTTTGCAACACATCA 2453
 2003 nAspProAlaAspProLysMetLeuGlnMetValLeuGlnGlySerValG 2020
 2454 GATCCCGGACAGACCCCAAAATGCTTCAGATGATGATCCAGGATTCGTAG 2503
 2020 LyrThrTrValAsnGlnGlyProLeuGluValAlaGlnValPheLeuSer 2036
 2504 GCACACAGTGAATCAAGCGGCTTTGGAAAGTGTCCAGGTTTTCGTCT 2553
 2037 GIuIleProSerAspProLysLeuPheArgHISAsnLysLeuArgLe 2053
 2554 GAATATCTATGTGACCCAAAGCTTTCACACATCAATTAATTAACCTGGACT 2603
 2053 uCysPheLysAspPheThrLysArgCysGIuAspAlaLeuArgLysAsnL 2070
 2604 CTGCTTTAAAGATTTTACTTAAAGGTGTGAGATGGCTTAAAGAAAAATA 2653
 2070 ySerLeuIleGlyProValGlnLysGluTrYGlnArgGluLeuGlyLys 2086
 2654 AGAGCTTAATTTGGCGGTTCAAAAGAGATCAAAAGGAATTTGCGGAAA 2703
 2087 LeuSerSerPro 2090
 2704 CTATCTTCGCGCT 2715
 seq_name: gb_pr:AK024436
 seq_documentation block:
 LOCUS AK024436 4577 bp mRNA linear PRI 29-SEP-2000
 DEFINITION Homo sapiens mRNA for FLJ00026 protein, partial cds.
 ACCESSION AK024436
 VERSION AK024436.1 GI:10440380
 KEYWORDS fis (full insert sequence).
 SOURCE Homo sapiens adult spleen cDNA to mRNA, clone:as00026.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 4577)
 AUTHORS Ohara,O., Nagase,T., Kikuno,R. and Okumura,K.
 TITLE The nucleotide sequence of a long cDNA clone isolated from human spleen
 JOURNAL Published Only in Database (2000) In press
 REFERENCE 2 (bases 1 to 4577)
 AUTHORS Ohara,O., Nagase,T., Kikuno,R. and Okumura,K.
 TITLE Direct Submission
 JOURNAL Submitted (24-AUG-2000) Osamu Ohara, Kazusa DNA Research Institute,
 Department of Human Gene Research, 1532-3, Yana, Kisarazu, Chiba
 292-0812, Japan (E-mail:cdna.info@kazusa.or.jp,
 URL:http://www.kazusa.or.jp/NEDO, Tel:81-438-52-3913,
 Fax:81-438-52-3914)

COMMENT

NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Kazusa DNA Research Institute.

FEATURES

Source

1. 4577
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="as00026"
/tissue_type="spleen"
/dev_stage="adult"
/note="vector:pluescriptII SK plus"
1. 3544
/gene="FLJ00026"
<1. 3544
/gene="FLJ00026"
/note="Start codon is not identified."
/product="FLJ00026 protein"
/protein_id="BAB15726.1"
/db_xref="GI:10440381"
/translation="KNIMSSKIDRNCRSMSTYCGSSSDAPSPAPRAPSKKHFE
LAIQVYVSTGVREIVFKYAMFEFFELIVSMAQHVHMDKRSFRTRSDRPMDDIT
TIVNVYSTGIALIVRPOKENEQAKMINSLAFYDLDSLMDRGVFNIRYKQSQ
SAKLSNLPILISMRLEFLRICSHHEYLNLNLFPMADAPSPSPSSISSONSSCS
FOOKIASMELDTSEYRQOHFLGTLFETLAALDDEGDISVOKRAVAISHILSS
HDLDPKRVKVKITAIYLPVLIIDLPOLCPEYADPRTYSGSDDEDEAG
AINONVALATAGNNPNKTSGLYSSIPYQYMLNADTRNMTFLMIMKADSL
IRKVIADLPSTQNLRIIDLFLICVLCEYKKGSSKSVSTQVQKSDVKARDEAL
RGGAREREMRRAPGNDPFGNENLRMKRQTHMRQAKELDKRAELDDEALISG
NLATEHLIILDMQENTIQASALDKDSLGLVILVLSLMDQSTYLLHCFEATL
RALIAKFGDLFEFEVPOCDFLCHOYLHSCSSMDYRSQACATLYILMFESGATN
FARKYQVMTSLASIVGARPDNEHLRSLRTILASEDTAMQTPPTOYEELIC
NUNSIIDTYKMEFOEDPMDIMLRIRAKTSQASPDRLITLQMAEKHTRKKRYT
EAMCVLHAALVAEYLSLMDHSTYLPVGSVFQNSNVLESVVSDEPLSDDEGV
CAGQYFESGELVGLDQAELEFSTGYETVNFVYKLVIPILAHKFRKLTLTHSKL
QRAFDSIVNKDRFMFGYFRVGFSGKFDLDEQFEVYEPATITLPEISHLEAFY
GOCGFAEFEVIRKDSPTVDKTLDPNKAITOIETFEVEPEYEMKDRVYFEENFIR
RPMYTPFTELEGRPRGLHEQRYRNLYLTMAHFPYIKRPSIVIOKEEPLTIEVAL
EDAKKTKLQALVAINDERPRAKMLQWLVGSSVATYNOGSELYAOUFLAETPDRPLY
RHHNKLILCFEFTMRGAEVKNKRLITADQREYQDELKNYKNKLNLRPMIEKRI
PELYKPLIFRSESQRDSFHRSSPRKCTQLDSGS"

CDS

BASE COUNT 1297 a 1037 c 1077 g 1166 t
ORIGIN

alignment_scores:

Quality: 3528.50 Length: 1144
Ratio: 3.634 Gaps: 11
Percent Similarity: 84.878 Percent Identity: 60.315

alignment_block:

US-09-737-246-2 x AK024436 ..

Align seg 1/1 to: AK024436 from: 1 to: 4577

```

951 SerThrGlnAlaMetAspArgSerCysAsnArgMetSerHisThrGln 967
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
14 TCTTCAAAAGATCGCGCATGCGCAACTGCGACGAAATGCTTACTATTGCG 63
967 uThrSerSerPheLeuGlnThr.....LeuThrGlyArgLeuProThrL 982
: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
64 TGGCAGTAGTAGTCCTCCAAAGTTCACTGACGCCCAAGCCAGCCAGCA 113
982 ySLysLeuPheHisGlnGluLeuAlaLeuGlnTrpValValCysSerGly 998
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
114 AAAAGCATTTCCATGAGAGCTTGCCTTCAGATGCTGTCAGCACCGGA 163
999 SerValArgIleSerAlaLeuGlnGlnAlaIleTrpPhePheGlnLeu 1015
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
164 ATGGTAGAGAAACAGCTTCAAGTATGCCGTGCTTCTTGGAGCTTCT 213

```

```

1015 tValLysSerMetValHisHisLeuTyrPheAsnAspLysLeuGlnAlaP 1032
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
214 GGTGAAGAGCATGCGCCAGACGCTACATTAACAGCAACAGGACAGATT 263
1032 roArgLysSerArgPheProGlnArgPheMetAspLysLeuAlaLeu 1048
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
264 TTGGAGAGCTCGTTTCTTGACCGCTTCAATGGATACATTAATCTATT 313
1049 ValSerThrIleAlaSerAspIleValSer.....ArgPheGlnLys 1062
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
314 GTTATATGTCATCCTCGCAATTTGCAGCCCTTTTATGTAACACAGAA 363
1062 sAspTrpGluMetValGlnArgLeuAsnThrSerLeuAlaPhePheLeu 1079
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
364 GGAATATGAACAGCGGAAAGATGAATCATCAGCTGCGCTTCTTCTTGT 413
1079 snAspLeuLeuSerValMetAspArgGlyPheValPheSerLeuLys 1095
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
414 ATGACCTTCTCTCCCTCATGATCGGGCTTGTGTAACTCATCATGAG 463
1096 SerCysTyrIleGlnValSerSerLysLeuTyrSerLeuProAsnPro 1112.
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
464 CATATTGCAAGCCAGCTGTCAGCCAACTCAGTAACCTTCCA..... 505
1112 tValLeuValSerLeuArgLeuAspPheLeuArgIleIleCysSerHis 1129
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
506 .ACGCTCATTTCCATGAGCTAGAGTTCCTGAGAATCCTCTGAGCAGAT 554
1129 LuHisTyrValThrLeuAsnLeu...ProCysSerLeuLeuThrPro 1144
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
555 AGCAATTCCTCAATCATGACCTTTTATGATGCTGATGCTGCCA 604
1145 AlaSerProSerProSerValSerSerAlaThrSerGlnSerSerGly 1161
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
605 ACATCTCCTCTGCTTCCATATCT.....TCCAGAACTCAAGCTC 645
1161 eSerThrAsnValGlnAspGlnLysIleAlaAsnMetPheGluLeuSer 1178
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
646 CTGCTCAGAGCTTCCAGAGCCAGAGATCGCAGATGTTGATCTGACT 695
1178 aLProPheArgGlnGlnHisTyrLeuAlaGlyLeuValLeuThrGln 1194
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
696 CCGAGTACCGGACACACACTCTCTCAACGCGCTCTCTTACACAGACT 745
1195 AlaValIleLeuAspProAspAlaGlnGlyLeuPheGlnLeuHisLys 1211
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
746 GCTGCTGCGCTGATGCCGAAGGGAAGATCAGCAAGATGACAAAGAA 795
1211 sValIleAsnMetValHisAsnLeuLeuSerSerHisAspSerAspPro 1228
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
796 AGCTGTACGTGCAATTCACAGCTGTAAAGTTCTCACGACCTGGACCC 845
1228 rGlyTyrSerAspProGlnIleLysAlaArgValAlaMetLeuTyrLeuPro 1244
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
846 GCTGTGTCAAACAGAGGTGAAGGTCAAAATCGCGCTTACCTTACCT 895
1245 LeuIleGlyIleIleMetGlnThrValProGlnLeuTyrAspPheThr 1261
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
896 TTAGTGTGCATCATTTTGGATGCTTGGCCACAGCTCTGTGACTTTAA 945
1261 uThrHisAsnGlnArgGlyArgProIleCysIleAlaThrAspAspTyr 1278
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
946 TGCAGATCTCGCAGATATCCGC.....ACCAGTGGCTCGGATGAG 986
1278 LuSerGluSerGlySerMetLysSerGlnThrValAlaMetAlaIleAla 1294
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
987 AACAGAAAGAGCGGTCCTTACCAACGAAATGTGCTGCGCATGACA 1036
1295 GlyThrSerValProGlnLeuThrArgProGlySerPheLeuLeuTrp 1311
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1037 GGAATTAATTTCAATTGAAAACA.....AGTGAATAGTGTGCTTCT 1080
1311 rThrSerGlyArgGlnHisThrThrPheSerAlaGluSerSerArgSerL 1328

```

```

1081 CTGCCCTATTAACGCTACACACTGCTGAACGGACACTACTCGCAACC 1130
      |||:||||:||||:||||:||||:||||:||||:||||:||||:
1328 euleuilecysleuLeuTrpValleuLysAsnIaAspLuhTrValleu 1344
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1131 TCATGATCTGCTCTCTGATCATGAAATATGCTGATCAGACCCCAT 1180
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1345 GlnLysTrpPheThrAspLeuSerValleuGlnLeuAsnArgLeuLeuAs 1361
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1181 AGGAAGTGGATTGCTGCTGCCATCAACGCAAGCTCAACAGATTTAGA 1230
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1361 pleuLeuTrpLeuCysValSerCysPheGluTrpLysGlyLysValP 1378
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1231 TCATCTTTTCATCTGTGTGTATGTTTGTATGAGTAAAGGAAACAGAG 1280
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1378 hegluArgMetAsnSerLeuThrPheLysSerLysSerLysMetArgIa 1394
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1281 CTGACAAATGCTACCTACCAAGTCTGCAAGACTCAAGGGATGTCAAG 1330
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1395 LysLeuGlnGluAlaIleLeuGlySerIleGlyAlaArgGlnMetVa 1411
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1331 CGGCTGGAAAGAGGCTTGCTCGTGGGGAAGGGCCAGAGGGAGATGAT 1380
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1411 ArgArgSerArgGlyGlnLeuGluArgSerProSerGlySerAlaPhe 1427
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1381 GCGCCGCC.....CGGCTCCAGGGAACGACGATTTTC 1412
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1428 ..GlySerGlnLysAsnLeuArgTrpArgLysAspMetThrIleTrpArg 1443
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1413 CAGGCTTAATGAAATTTGAGATGGAAAGAAAGACACACATTTGGCG 1462
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1444 GlnAsnThrGluLysLeuAspLysSerArgAlaGluIleGlnHisGluAl 1460
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1463 CAAGCTAATAGAGAGCTAGATAAACAAGCGGAGTTAGATCAGAAAGC 1512
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1460 AleuIleAspGlyAsnLeuAlaThrGluAlaAsnLeuIleIleuAspTr 1477
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1513 CTGTGATCACTGGCAATCTGGCTACAGAAACACATTTAATCATCCCTG 1562
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1477 hrLeuGluTrpValleuValGlnThrValSerValThrGluSerLysL 1493
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1563 TCCAGGAAACATTTATCCAGGCGGAGCTGGCTGTGAGCTGTAAGACAG 1612
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1494 IleLeuGlyGlyValleuLysValleuLeuHisSerMetAlaCysAsnG 1510
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1613 CTGCTGGAGGTGTCTGTGAGGCTGCTGTGATCTCTGAACTGTGATCA 1662
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1510 nSerAlaValTrpLeuGlnHisCysPheAlaThrGlnArgAlaLeuVal 1527
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1663 GAGTACCACTACCTGACTCACTGCTTGCACAACTCCGTCTCATCG 1712
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1527 erLysPheProGluLeuLeuPheGluGluGluTrpGluGlnCysLasp 1543
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1713 CCAAGTTTGGAGACTTACTCTTCGAAAGAGAGGTGGAAAGGTTTCGAC 1762
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1544 LeuCysLeuArgLeuLeuArgHisCysSerSerSerIleGlyThrIleAr 1560
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1763 CTATGTACCAAGTCTCGACACATGCAGACAGCATGATGTCAACCG 1812
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1560 gSerHisProSerAlaSerLeuTrpLeuLeuMetArgGlnAsnPheGlu 1577
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1813 GAGCCAGACCTGTCACCCCTTTACCTCTCATGAGTTTCAGTTTGGAG 1862
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1577 legLysAsnAspPheAlaArgValLysMetGlnValProMetSerLeu 1593
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1863 CCACCAAGTATTTTGCAGAGATGAAGATGCAGATACCATGTCCCTGGCA 1912
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1594 SerLeuValGlyThrSerGlnAsnPheAsnGluGluPheLeuArgArg 1610
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1913 TCTTTGTGGGAAGACACAGACTTTAATGAAGAGCACCTGAGAAGATC 1962
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1610 rLeuLysThrIleLeuThrTrpAlaGluGluAspLeuGluLeuArgGlu 1627
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

1663 CTGAGAGCAATTTTGGCTATTCTCAGAAAGACACAGCCATGCAGATGA 2012
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1627 hrThrPheProAspGlnValGlnAspLeuValPheAsnLeuHisMetIle 1643
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2013 CTCCTTTTCCACCAGGTGGAGAACTCTCTTAATCTGAATAGATC 2062
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1644 LeuSerAspThrValLysMetLysGlnHisGlnGluAspProGluMetLe 1660
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2063 TTATATGACACAGTGAATATAGGAATTTTCAGAAATCTGATGATGCT 2112
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1660 nIleAspLeuMetTrpArgIleAlaLysGlyTrpGlnThrSerProGlu 1676
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2113 TATGATCTCATGTACAGAAATTTGCCAAGAGTTTACAGCATCTCCGTG 2162
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1677 ..ArgLeuThrTrpLeuGlnAsnMetAlaGlyLysHisSerGluArgSer 1692
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2163 TCGGCTGACCTGGCTCCAGAACTGGCAAGAAACACACAGAAAGAG 2212
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1693 AsnHisAlaGluAlaAlaGlnCysLeuValHisSerAlaIleuValAl 1709
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2213 TGCTACACGGAGGCTGCCATGTGCTGTGCACGCCCTCGTTAGTGGC 2262
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1709 aGluTrpLeuSerMetLeuGluAspArgLysTrpLeuProValGlyCysV 1726
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2263 TGAGTATCTGAGCATGTCTGGAGACACACACTACCTCCGTGGCAGTG 2312
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1726 aIleThrPheGlnAsnIleSerSerAsnValleuGluGluSerAlaValSer 1742
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2313 TCACCTTCCAGAAATATTTCTTCCAAATGTCTGAGAGGTGTGTGCTCT 2362
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1743 AspAspValAlaSerProAspGluGluGlyLysSerGlyLysTrpArg 1759
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2363 GAGAGACACCCCTGTACCTGCACAGAGATGGGTGTGGCCAGCCAGTACT 2412
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1759 eThrGluSerGlyLeuValGlyLeuLeuGluGlnAlaAlaIleSerPhe 1776
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2413 CACCGAGAGTGGCTGTAGGCTCTCGAGACAGCGCCGAGACTCTTCA 2462
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1776 erMetAlaGlyMetTrpGluAlaValAsnGluValTrpLysValIle 1792
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2463 GCACGGAGGCTTATATGAGACAGTTAATAGGCTTCAACAGTGTGATC 2512
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1793 ProIleHisGluAlaAsnArgAspAlaLysLysLeuSerThrIleHisG 1809
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2513 CCCATCTTAAGAAGGCATCGAAGATTTCCGAACTGACATCTACACAG 2562
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1809 LysLeuGlnGluAlaPheSerLysIleValHisGlnSerThrGlyTrpG 1826
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2563 CAAGCTCAAGAGACCTTCAGACAGCATCGTTACAAGAT.....CATA 2606
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1826 LuArgMetPheGlyThrTrpPheArgValGlyPheTrpGlyThrLysPhe 1842
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2607 AGAAGATGTTTGAACCTACTTCCGAGTTGTTCTTGGATCCAATTT 2656
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1843 GlyAspLeuAspGluGlnGluPheValTrpLysGluProAlaIleThrL 1859
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2657 GGGGATTTGGATGACAGCAATTTGTCTACAAAGACCTCAATTTACCA 2706
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1859 sLeuAlaGluIleSerHisArgLeuGluGlyPheTrpGlyGluArgPhe 1876
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2707 GCTTCCGAGATCTCAATAGACTAGAGCATTTATGTCATGTTTGG 2756
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1876 LysGluAspValValGluValIleLysAspSerAsnProValAspLysC 1892
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2757 GTGCAGATTTTGTGAAGTGAATTAAGACTCCCTCTGTGGCAAAAC 2806
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1893 LysLeuAspProAsnLysAlaTrpIleGlnIleThrTrpValGluProTy 1909
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2807 AAGTGGATCTTAACAAGGCTACATACAGATCTTTGTGTGAGCCCTA 2856
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1909 rPheAspThrTrpGluMetLysAspArgIleThrTrpPheAspLysAsn 1926
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2857 CTTTGATGATGATGATGATGAAGACAGGCTCACATCTTGAAGAAAT 2906
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```
1926 yfAsnleuArgrPheMetYrCysThrProPheThrLeuAspGlyArg 1942
:::|||||
2907 TCACCTCCGGGGTTTCATGTACACACCCCGTTCACCCGTGGAGGGCGG 2956
|||||
1943 AlaHisGlyGluLeuHisGluGlnPheLysArgLysThrIleLeuThrTh 1959
:::|||||
2957 CCTCGGGAGAGCTGCATGACGACGACGAGACACACAGCTCCGACCAC 3006
rserHisAlaPheProTyrIleLysThrArgValAsnValThrHisLysG 1976
|||||
3007 TATGCACCGCTTCCCTACATCAACAGACGATCAGCGTCAATCCAGAAAG 3056
1976 IuGluIleIleLeuThrProIleGluValAlaIleGluAspMetGlnLys 1992
|||||
3057 AGGAGTTTGTTTGACACCGATGTGAAGTTGCCATGTGAAGACATGAGAAG 3106
1993 LysThrGlnGluLeuAlaPheAlaThrHisGlnAspProAlaAspProLys 2009
|||||
3107 AAGACCCCTGCACTTAGCAGTTGCCATTAACGAGAGCCCGCTGATGCAG 3156
2009 smetLeuGlnMetValLeuGlnGlySerValGlyThrThrValAsnGlnG 2026
|||||
3157 GATGCTTCAGATGCTGCTGCAAGGCTGTGGGAGCTACTGTAATCAGG 3206
2026 lYProLeuGluValAlaGlnValPheLeuSerGluIleProSerAspPro 2042
|||||
3207 GACCACTGAGAGTACCCCAAGTGTGCTGGAATTCCTGCTGATCCA 3256
2043 LysLeuPheArgHisHisAsnLysLeuArgLeuLysPheLysAspPheTh 2059
|||||
3257 AAACCTATCGACATCACACAAAGTTGAGGTTATGCTTTAAGAAATTCAT 3306
2059 rLysArgCysGluAspAlaLeuArgLysAsnLysSerLeuIleGlyProV 2076
|||||
3307 CATGAGATGTGTGAAGCTGTAGAGAAAAACAAGCGTCTCATACAGGCAG 3356
2076 aIGlnLysGluTyrGlnArgGluLeuGlyLys 2086
|||||
3357 ACCAGAGGGAATATCAGCAGAACTCAAAAG 3388
```


50 Sequence 6828 BP; 2082 A; 1384 C; 1432 G; 1930 T; 0 other;

alignment_scores:

Quality: 10815.00 Length: 2090
Ratio: 5.175 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-737-246-2 x AAH41934 ..

Align seg 1/1 to: AAH41934 from: 1 to: 6828

```
1 MetAlaGluArgArGAlaPheAlaGlnLysIleSerArgThrValAlaL 17
|||||
22 ATGGCGGAGCGCCGCCCTTCGCCGAGAGATCAGCAGACGGTGGCAGC 71
|||||
17 aglValArgLysGlnIleSerGlyInTySerGlySerProGlnLeuL 34
|||||
72 CGAAGTTAGGAAGCAGATCTCCGACAATATAGTGTCTCCCAACTGC 121
|||||
34 eulYsAsnLeuAsnIleValIglYAsnIleSerHisThrThyAlPro 50
|||||
122 TCAAAAACCTTAATATGTGGCAATATATCCATCACACCAAGTGGCC 171
|||||
51 LeuThrGluAlaValAspProValAspLeuGluAspTyrLeuIleThrHi 67
|||||
172 CTTACCGAAGCAGTAGATCCAGTGGATTGGAGAATTACCTATTAATCA 221
|||||
67 sProLeuAlaValAspSerGlyProLeuArgAspLeuIleGluPhePro 84
|||||
222 TCCTTGGCTGTGATTTGGGCGCTTACGGGATTTGATTGAATTTCCGC 271
|||||
84 roAspAspIleGluValValIYrSerProArgAspCysArgThrLeuVal 100
|||||
272 CAGATGATATGAAAGTGTATAGTCTCGGGAGCTGCAGAACTCTGTGT 321
|||||
101 SerAlaValProGluGlnSerGlnMetAspProHisValArgAspCysI 117
|||||
322 TCAGCTGTACCTGMAAGAAAGTGAATGATCCACATGTTAGAGACTGAT 371
|||||
117 earGSerTyrThrGluAspTrpAlaIleValIleArgLysTyrHisLysL 134
|||||
372 AAGAAGTTATACAGAAGACTGGGCAATTGTCTACAGAAAATATCATTAAT 421
|||||
134 euGlyThrGlyPheAsnProAsnThrLeuAspLysGlnLysGlnArgGln 150
|||||
422 TGGGAACAGGATTTATCCCAATACATTAGATAAACAGAAAAGAGGCAA 471
|||||
151 LysGlyLeuProLysGlnValPheGlnSerAspGluAlaProAspGlyAs 167
|||||
472 AAAGGTTTCCAAACAAACAGTTTGTGATCTGATGAAGCTCCGAGATGCCAA 521
|||||
167 nSerTyrGlnAspAspGlnAspAspLeuLysArgArgSerMetSerIleA 184
|||||
522 CAGCTACCGAGATGATCAAGATGACCTTAAAGAGCGTTCAATGTCAAATAG 571
|||||
184 spAspThrProArgGlySerTrpAlaCysSerIlePheAspLeuLysAsn 200
|||||
572 ATGATATCCCAAGGGGTAGCTGGCGCTGTAGTATCTTGACTTGAAAATAT 621
|||||
201 SerLeuProAspAlaLeuLeuProAsnLeuLeuAspArgThrProAsnG 217
|||||
622 TCACCTTCCTGATGCTTGTGCTCCCAATTTACTGTGATCGAACTCCCAATGA 671
|||||
217 uGluIleAspArgGlnAsnAspAspGlnArgLysSerAsnArgHisLysG 234
|||||
672 AGAAATAGACCGTCAGAATGATGACCAAAAGAAATCAAAACGCTCACAAG 721
|||||
234 IuLeuPheAlaLeuHisProSerProAspGluGluGluProIleGluArg 250
|||||
722 AACCTTTTGGCTTGTGATCCATCACCAGATGAGAGAAACCAATAGACGG 771
|||||
```

```
251 LeuSerValProAspIleProLysGlnHisPheGlyGlnArgLeuVal 267
|||||
772 CTTAGTGTCTGATATATCCAAAGAAACATTTGTGTAAGAACTCTGTGT 821
|||||
267 LysCysLeuSerLeuLysPheGluIleGluIleGluProIlePheAla 284
|||||
822 AAATGCTTATCACTCAAGTTGAAATTTGAAATTTGAACCCATTTTGGCAA 871
|||||
284 erLeuAlaLeuTyrAspValLysGluLysLysLysIleSerGluAsnPh 300
|||||
872 GTTGGCTTTATATGATGATCAGAGAAAAGAAAAGATTTTCAGAAAACCTTT 921
|||||
301 TyrPheAspLeuAsnSerGluGlnMetLysGlyLeuLeuArgProHisVa 317
|||||
922 TATTTGACCTTAATTTCTGACAGATCAAAAGGTTGTACGCTCACATGT 971
|||||
317 LProProAlaAlaIleThrThrLeuAlaArgSerAlaIlePheSerIleT 334
|||||
972 ACCACCTGTGCTCCATTAATACCTGCGCAAGATCAGCAATTTTTCATCA 1021
|||||
334 hrTyrProSerGlnAspValPheLeuValIleLysLeuGluLysValLeu 350
|||||
1022 CTATTCCTTCCCAAGATGTTTCTGTGTAATTAAGCTAAGAAAAGTCTTA 1071
|||||
351 GlnGlnGlyAspIleGlyLysCysAlaGluProThrMetIlePheLysG 367
|||||
1072 CAGCAAGAGACATTTGGAGAGTGTGCAGAACCATATATGATTTTCAAGA 1121
|||||
367 uAlaAspAlaThrLysAsnLysGluLysLeuLysLysSerGln 384
|||||
1122 AGCAGATGCCCAACAAGATTAAGAAAACCTGAGAAAACCTGAGAGTCAAG 1171
|||||
384 laAspGlnPheCysGlnArgLeuGlyLysTyrArgMetProPheAlaTrp 400
|||||
1172 CAGATCAGTTTTGGCCAAAGACTTGGGAAATATCGCAATCCCTTTGTCTGG 1221
|||||
401 ThrAlaIleHisLeuMetAsnIleValSerSerAlaGlySerLeuGluArg 417
|||||
1222 ACTGCATCCATTTTAATGAATATTTGTTAGCAGTGTCTGGAGTTTGGAAAG 1271
|||||
417 gAspSerThrGluValGluIleSerThrGlyLysArgLysGlySerTrp 434
|||||
1272 AGATTCTAACAAGAACTAGAAATCACTACTGAGAGAACGAAAGGCTGTGCT 1321
|||||
434 erGluArgArgAsnSerSerIleValIglYArgArgSerLeuGluArgThr 450
|||||
1322 CAGAGAGAGGAATTTCTATGATTTGTGGCAGACGATCATTGGAAGGACA 1371
|||||
451 ThrSerGlyAspAspAlaCysAsnLeuThrSerPheArgProAlaThrLe 467
|||||
1372 ACAAGTGGAGATGATGCTTGTACTTGACGACGCTTTCGACACACTACTCT 1421
|||||
467 uThrValThrAsnPhePheLysGlnGluLysAspArgLeuSerAspGluA 484
|||||
1422 CACAGTGACAAATTTTAAAGCAGAGAGAGACCGCTTAAGTGAATGAAG 1471
|||||
484 splLeuTyrLysPheLeuAlaAspMetArgArgProSerSerValLeuArg 500
|||||
1472 ATCTCTTCAAAATTTCTTGTGATATGAGAGGCCATCTTGTCTTACGG 1521
|||||
501 ArgLeuArgProIleThrAlaGlnLeuLysIleAspIleSerProAlaPr 517
|||||
1522 CGACTAAGACCTATTTACAGCTCACTCAAAATAGACATTTCTCCCGCAC 1571
|||||
517 oGluAsnProHisTyrCysLeuThrProGluLeuGlnValLysLeu 534
|||||
1572 TGAATAATCCCATATATGCTTAACCTCGAGAGCTGCTTCAAGTGAACCTTT 1621
|||||
534 yrProAspSerArgValArgProThrArgGluIleLeuGluPheProAla 550
|||||
1622 ACCCTGACAGTATGAGTTAGACTTACCAAGAGAAATCTTAAAGTTTCCCGCA 1671
|||||
551 ArgAspValTyrValProAsnThrThrTyrArgAsnLeuLeuTyrIleTy 567
|||||
```

```

|||||
1672 AGGAGATTATGTTCCAAACACTACTTACAGAAATCTTCTACATATA 1721
567 rProGlnSerLeuAsnPheAlaAsnArgLInGlySerAlaArgAsnIleT 584
|||||
1722 CCTCAGAGTCTTAATTTTGGCAATGTCAGAGTTCTGCTAGAAATATATA 1771
584 hValLysValGlnPheMetYrGlyGluAspProSerAsnAlaMetPro 600
|||||
1772 CAGTGAAGTCCAGTTTATGTATGAGAGATCCAAAGCAATGCCATGCCG 1821
601 ValLlePheGlyLysSerSerGlyPheSerLysGluAlaIleTyrTh 617
|||||
1822 GTAATCTTGGTAAATCTAGCTGTGAGATTTTCAAGAGAACCTATAC 1871
617 rAlaValAlaIYrHisAsnArgSerProAspPheHisGluIleLysV 634
|||||
1872 AGCCGATGATATCATACAGGCTCTGATTTTCTAGAGAAATCAGAG 1921
634 aLysLeuProAlaThrLeuThrAspHisHisLeuLeuPheThrPhe 650
1922 TTAAAGCTCTGCTACTTAACTGACCATCATCACTGCTTTTACTTTT 1971
651 TyrHisValSerCysGlnGlnLysGlnAsnThrProLeuGluThrProVa 667
|||||
1972 TATCATGTAGTGTGTCAACAAAACAAATATCTCTTGAACACCACT 2021
667 LgLYrYrThrTrpLleProMetLeuGlnAsnGlyArgLeuLysThrGlyS 684
|||||
2022 TGGATATACATGGATACCAATGCTTCAGATGACGTTGAAAGCTGGCC 2071
684 lPheCysLeuProValSerLeuGluLysProGlnAlaTyrSerVal 700
|||||
2072 AGTTTGGCTTGGCCAGTCTCATTTGGAATAAACCAACAGGCTTATCTGTA 2121
701 LeuSerProGluValProLeuProGlyMetLysTrpValAspAsnHisLys 717
|||||
2122 CTGTCTTCGAGAGTCTCTACTCTGACATGAATGGGTAGATATCTACAA 2171
717 sGlyValAlaPheAsnValGluValAlaValAlaValSerSerIleHisThrGlnA 734
|||||
2172 AGGTGTTTTTAATGTTGAAGTGTCTGTTTCTGCTATCATCACAGAG 2221
734 sPProTyrLeuAspLysPhePheAlaLeuValAsnAlaLeuAspGluHis 750
|||||
2222 ATCCTTATCTTGACAAATTTTGGTCTGTCATGCTCTGGATGACAC 2271
751 LeuPheProValArgIleGlyAspMetArgIleMetGluAsnAsnLeuG 767
|||||
2272 CTGTCTCCAGTCCGAATGGGACATGGCAATCATGGAATAATACTTAGA 2321
767 uAsnGluLeuLysSerSerIleSerAlaLeuAsnSerSerGlnLeuGluP 784
|||||
2322 AAATGAAATGGAAAGCATATTTTCAGCACTGAATTCATCCACAGCTGGAAC 2371
784 roValValAlaArgPheLeuHisLeuLeuLeuAspLysLeuIleLeuLeuVal 800
|||||
2372 CAGTGGTCCGATTTCTTCATCTCTCTGATGATTAACCTGATACTTAACT 2421
801 lLeaArgProProValIleAlaGlyGlnIleValAsnLeuGlyGlnAlaSe 817
|||||
2422 ATTAGACCTCTCTGCTATGCTGGCCAAATAGTTAACTAGTCAACATAC 2471
817 rPheGluAlaMetAlaSerIleLeuAsnArgLeuHisLysAsnLeuGlnG 834
2472 TTTTGAAGCCATGCAATCAATTAATAATGACTTCAACAAAACCTTGGAGG 2521
834 lYAsnHisAspGlnHisGlyArgAsnSerLeuLeuAlaSerTyrIleHis 850
|||||
2522 GAATCATGTACAGCATGGCAGAAACAGCTCTTGCATCATATATATCAT 2571
851 TyrValPheArgLeuProAsnThrTyrProAsnSerSerSerProGlyP 867
|||||
2572 TATGTTTTCCGCTACCAAAATACTTACCTAATTCATCATCACAGGCTCC 2621
867 oGlyGlyLeuGlyLysSerValHisTyrAlaThrMetAlaArgSerAlaV 884
|||||
2622 TGGGGGTTTGGAGAGATCAGTGCATTTATGCCACATGCGTAGATCTGCCG 2671
884 aLArgProAlaSerLeuAsnLeuAsnArgSerArgSerLeuSerAsnSer 900
|||||
2672 TGAGACCTGCAGACCTTAATTTAAATGTTCTCGAACCTTAGTAATAGC 2721
901 AsnProAspLysSerGlyThrProThrSerProAspArgLysValAlaArgSe 917
|||||
2722 AATCCAGATATATCTGGGACTCCACGTCACACAGATATATGAAGTTGATC 2771
917 rIleIleGlySerLysGlyLeuAspArgSerAsnSerTrpValAsnThrG 934
2772 AATCATCGGAGTAGTAGGTTTAGATCGCTCCAAATCTCTGGGTTAAACCTG 2821
934 lYgLYrProLysAlaAlaProTrpGlySerAsnProSerProSerAlaGlu 950
2822 GTGGTCCAAAAGCTGCCCATGGGATCCAAACCCAGTCCAAAGTGCAGAA 2871
951 SerThrGlnAlaMetAspArgSerCysAsnArgMetSerSerHisThrG 967
|||||
2872 TCAACACAGGCTATGAGATCGAAGTTGTAATCGTATGCTTGCACACAGA 2921
967 uThrSerSerPheLeuGlnThrLeuThrGlyArgLeuProThrLysLysL 984
|||||
2922 GACGTCAGAGTTTCTTACAAACATTTAACGGAGCGCTTACCAACTAAAGC 2971
984 eupheHisGluGluLeuAlaLeuGlnThrPValValCysSerGlySerVal 1000
|||||
2972 TTTTTCACAGAGAGACTGGCTTGGCAGTGGGTTGTTTCAGTGCACACGCTT 3021
1001 ArgGlnSerAlaLeuGlnGlnAlaThrPhePheGluLeuMetValLy 1017
|||||
3022 CGGGATACGCTTTGGCAACAGCCGTGTTCTTTTGAATTAATGATGTA 3071
1017 sSerMetValHisHisLeuTyrPheAsnAspLysLeuGluAlaProArgL 1034
|||||
3072 GAGCATGCTGCACCATTTATCTTAATGATTAACCTTGAGGCTCCAAAGCA 3121
1034 ySSerArgPheProGluArgPheMetAspAspIleAlaAlaLeuValSer 1050
|||||
3122 AAAGCGTTTTCCAGAACGTTTTCATGATGATGACATTCAGCTCTGTGACG 3171
1051 ThrIleAlaSerAspIleValSerArgPheGlnLysAspThrGluMetVa 1067
|||||
3172 ACGATTGCTAGTATATATGTTTCACGATTTCCAGAAAGCACAGAAATGT 3221
1067 lGluArgLeuAsnThrSerLeuAlaPhePheLeuAsnAspLeuLeuSerV 1084
|||||
3222 TGAGAGACTCAATACAGCCTTGCAATTTCTTCATATGATCTGTGTCTG 3271
1084 aLMetAspArgGlyPheValPheSerLeuIleLysSerCysTyrLysGln 1100
|||||
3272 TTATGACAGAGAGATTTGTTTATAGCCTTATTAAGTCCCTGCTATTAACAG 3321
1101 ValSerSerLysLeuTyrSerLeuProAsnProSerValLeuValSerLe 1117
|||||
3322 GTGTCTTAAAGCTTTACTCATTTACCGAATCCAGTGTCTGGGTGCTT 3371
1117 uArgLeuAspPheLeuArgIleIleCysSerHisGluHisTyrValThrL 1134
|||||
3372 GAGGCTGATTTTCTACGAATCATCTGAGTCATGAGACATATTTTACAT 3421
1134 euAsnLeuProCysSerLeuLeuThrProProAlaSerProSerProSer 1150
|||||
3422 TAAACTTACCCGTCAGCTTACTACTCCACCTGCATCTCCATCACCTTCT 3471
1151 ValSerSerAlaThrSerGlnSerSerGlyPheSerThrAsnValGlnAs 1167
|||||
3472 GTTCTCTGCAACATCTCAGAGTTCTGATTTTCTAGGAATGTACAGA 3521

```

1167 pGlnuysIleAlaAsnMetPheGluLeuSerValProPheArgGlnGln 1184
 |||||
 3522 CCAAAAGATGGCAATATCTTTGAATATCCGTGCTTCCGCCAACAGC 3571
 |||||
 1184 IsTyrLeuAlaGlyLeuValLeuThrGluLeuAlaValIleLeuAspPro 1200
 |||||
 3572 ATTATTTGGCAGGACTGTGTAAACAGAGCTGCTGATTTTACACCT 3621
 |||||
 1201 AspAlaGluGlyLeuPheGlyLeuHisLysValIleAsnMetValHi 1217
 |||||
 3622 GATGCGAAGGACTGTTGGATTGCATAGAAAGTCATCAATATGGTCA 3671
 |||||
 1217 sAsnLeuSerSerHisAspSerAspProArgTyrSerAspProGlnI 1234
 |||||
 3672 CAATTTACTCTCCAGTCACGACTCAGACCCGGGTACTCTGACCTCGA 3721
 |||||
 1234 IeLysAlaArgValAlaMetLeuTyrLeuProLeuIleGlyIleIleMet 1250
 |||||
 3722 TAAAGGCTCGAGTGGCCATGTTGTATCACTGATTTGATATCATG 3771
 |||||
 1251 GluThrValProGlnLeuTyrAspPheThrGluThrHisAsnGlnArgG 1267
 |||||
 3772 GAACCTGACCTCAGCTGTATGATTTTACAGAACTCAACATCAACGAG 3821
 |||||
 1267 YArgProIleCysIleAlaThrAspPtyrGluSerGluSerGlySerM 1284
 |||||
 3822 AAGACCAATTTGTATAGCCACATGATGATTGAAGTCAAGACCGGAATA 3871
 |||||
 1284 eTilSerGlnThrValAlaMetAlaIleAlaGlyThrSerValProGln 1300
 |||||
 3872 TGATTAAGCCAGACCGTGGCATGCAATCGCAGGACATCGTCCCTCAA 3921
 |||||
 1301 LeuThrArgProGlySerPheLeuLeuThrSerThrSerGlyArgGlnHi 1317
 |||||
 3922 CTAAACAAGCCCTGGCAAGTTTCTCTCAAGTCACAGATGGCAGGACA 3971
 |||||
 1317 sThrThrPheSerAlaGluSerSerArgSerLeuLeuIleCysLeuLeuT 1334
 |||||
 3972 CACTACCTTTTCAGCAGAAATCAAGTCGAAGCCTTTGATCTGTCTACTT 4021
 |||||
 1334 rValLeuLysAsnAlaAspGluThrValLeuGlnLysThrPheThrAsp 1350
 |||||
 4022 GGGTTCTCAAAATGCAGATGAACAGTTCTACAGAAAGTGGTTACAGAT 4071
 |||||
 1351 LeuSerValIleuGlnLeuAsnArgLeuAspLeuLeuTyrLeuCysVal 1367
 |||||
 4072 CTCTCAGCTTTCAGACTTAACCGGCTATTAGTCTGCTTATCTCTGTGT 4121
 |||||
 1367 IserCysPheGluTyrLysGlyLysValPheGluArgMetAsnSerL 1384
 |||||
 4122 GTCTTGCTTTGAGTATAAGGAAAAAAGTGTGGTGAACGAAATGAACTACT 4171
 |||||
 1384 eutThrPheLysLysSerLysAspMetAlaGlyLysLeuGluGluAlaIle 1400
 |||||
 4172 TGACCTTTTAAGAAATCAAAAGCATGAGCAAAAGCTTGAAGAAGCTATTT 4221
 |||||
 1401 LeuGlySerIleGlyAlaArgGlnGluMetValArgArgSerArgGlyG 1417
 |||||
 4222 CTGGGAGACATAGTCCAGGCAAGAAATGCTACGGCCAGAACCCGAGACA 4271
 |||||
 1417 nLeuGluArgSerProSerGlySerAlaPheGlySerGlnGluAsnLeuA 1434
 |||||
 4272 GCTCGAAGAAAGCCATCTGGAAGTGCTTTGGAAGTCAGAAGAAATTTGA 4321
 |||||
 1434 rGTTrArgLysAspMetThrHisTTPArgGlnAsnThrGluLysLeuAsp 1450
 |||||
 4322 GGTGAGGAAGAATATGTACTCAGTGGCTCAAAACACAGAGAACTTGAC 4371
 |||||
 1451 LysSerArgAlaGluIleGluHisGluAlaLeuIleAspGlyAsnLeuAl 1467
 |||||
 4372 AAATCAAGACGAGATTTGAACACGAAAGCACTGATTTGATGAAACCTGGC 4421

1467 aThrGluAlaAsnLeuIleIleLeuAspThrLeuGluIleValIValGlnT 1484
 |||||
 4422 TACGAAGCAAACTTATCTATTTAGATACATTAGAGATTGTGTTCACA 4471
 |||||
 1484 hrValSerValThrGluSerLysGluSerIleLeuGlyGlyValLeuLys 1500
 |||||
 4472 CCGTTTCTGTAAAGGAATCCAAAGAGAGCATTCCTGGTGGAGTGTAAAA 4521
 |||||
 1501 ValLeuLeuHisSerMetAlaCysAsnGlnSerAlaValTyrLeuGlnHi 1517
 |||||
 4522 GTGCTACTACACAGCATGGCTGTAAACCAAGTCAAGTTTATCTACACA 4571
 |||||
 1517 sCysPheAlaThrGlnArgAlaLeuValSerLysPheProGluLeuLeuP 1534
 |||||
 4572 CTGTTTTCCTACACAGAGACCTTGTTCAAGTTTCCGAACTCTTAT 4621
 |||||
 1534 heGluGluGluThrGluGlnCysAlaAspLeuCysLeuArgLeuLeuArg 1550
 |||||
 4622 TTGAAAGAGAGACAGACAGTGTGTGATTTATGCCACAGCTTTCGGA 4671
 |||||
 1551 HisCysSerSerSerIleGlyThrIleArgSerHisProSerAlaSerie 1567
 |||||
 4672 CACTGTAGCAGTAGCATGGTACAAATAGGTCAACCCAGTCCCTCCCT 4721
 |||||
 1567 uTyrLeuLeuMetArgGlnAsnPheGluIleGlyAsnAsnPheAlaArgV 1584
 |||||
 4722 TTACCTTCTATAGAGCAAACTTTGAGATTGGGAATTACTTTGCCAGGG 4771
 |||||
 1584 aLysMetGlnValPrometSerLeuSerSerLeuValGlyThrSerGln 1600
 |||||
 4772 TTAAATAGCAGTACCAATGATCTCATTCCTTGGGGGACATCTCGAG 4821
 |||||
 1601 AsnPheAsnGluGluPheLeuArgArgSerLeuLysThrIleLeuThrTy 1617
 |||||
 4822 AATTTTATGAAGAATTTTAAGACGTTCTTAACACATATATTTGACATA 4871
 |||||
 1617 rAlaGluGluAspLeuGluLeuValArgGluThrThrPheProAspGlnValG 1634
 |||||
 4872 TGCGAAGAAAGATCTGGAATTTGAGGGAACACATTTCCGATCAGGTGC 4921
 |||||
 1634 IAspLeuValPheAsnLeuHisMetIleLeuSerAspThrValLysMet 1650
 |||||
 4922 AGCATCGGTTTCAATCTCCATATGATCTTTCTGATCTGGAAGAAAG 4971
 |||||
 1651 LysGlnHisGlnGluAspProGluMetLeuIleAspLeuMetTyrTrgI 1667
 |||||
 4972 AAGGAACACGAGAGATCCTGAATGTGATTGATTTAATGTATACAGAT 5021
 |||||
 1667 eAlaLysGlyTyrGlnThrSerProGluArgLeuThrTTPLeuGlnAsnM 1684
 |||||
 5022 TGCCAAGGGTTACCAGACCTCTCCAGAGGATTTGACCTGGTTGCCAACA 5071
 |||||
 1684 eTAlaGlyLysHisSerGluArgSerAsnHisAlaGluAlaIleGlnCys 1700
 |||||
 5072 TGGCAGGCAAGCACTCAGAAAGCAATCATGCTCAAGCTGCACAGTGT 5121
 |||||
 1701 LeuValHisSerAlaIleLeuValAlaGluTyrLeuSerMetLeuGlnLys 1717
 |||||
 5122 CTAGTCCACACACAGACACTTGTGCTGATTTATTTAGCATGTCTGGAGGA 5171
 |||||
 1717 ParGlyLysTyrLeuProValGlyCysValThrPheGlnAsnIleSerSera 1734
 |||||
 5172 CCGGAATATCTCTGTTGGGATGTGAACATTTGAGATATTTATCTCA 5221
 |||||
 1734 snValLeuGluGluSerAlaValSerAspAspValValSerProAspGlu 1750
 |||||
 5222 ATGTTTGAAGAATCTGGCGTCTCAGATGATGTGTATCTCCAGATGGA 5271
 |||||
 1751 GluGlyIleCysSerGlyLysTyrPheThrGluSerGlyLeuValGlyLe 1767
 |||||
 5272 GAAGGATTCGCTCTGAAATATCTTTACTGATCAGGACTGTGGGATTT 5321
 |||||
 1767 uLeuGluGlnAlaIleAlaSerPheSerMetAlaGlyMetTyrGluAlaV 1784


```

|||||
5322 ACTGAACAGACGCTGCTCCTCTCATGCGTCATGATGACAG 5371
1784 aIAsnGIValTyrTyrValLeuIleProIleHisGLuAlaAsnArgAsp 1800
|||||
5372 TTAATGAAGTTACAAAGTACTTATCTTATTCATGAAAGCTATACGGGAT 5421
1801 AlAlsLysLeuSerThrIleHisGLysLeuGlnGluAlaPheSerLys 1817
|||||
5422 GCAAGGAACATCCACATTCATGCTAACTCAAGAAGCATTCAGCAA 5471
1817 sIleValHisGlnSerThrGlyTrpGluArgMetPheGlyThrTyrPhea 1834
|||||
5472 AATGTTCATCAGACTGCTGCGTGGAGCGGATGTTGGACCTATTTC 5521
1834 rGValGlyPheTyrGlyThrLysPheGlyAspLeuAspGluGlnGluPhe 1850
|||||
5522 GGTGTGGTTTATGTAACCAAGTTGCGGATTTGATGACAAAGAAATT 5571
1851 ValTyrLysGluProAlaIleThrLysLeuAlaGluIleSerHisArgLe 1867
5572 GTTTCACAGAGCGCTGCATTAACCAACTGACAGATATCTCCACAGATT 5621
1867 uGluGlyPheTyrGlyLysArgPheGlyGluAspValValGluValIleL 1884
|||||
5622 GGAGGATTTTACGGGAAGATTGAGAGGATGTTGTTGAAGTAAACA 5671
1884 yAspSerAsnProValAspLysCysLysLeuAspProAsnLysAlaTyr 1900
|||||
5672 AAGACTTATATCTGTAGACAAAGTAAATTAGATCTTACACAGCATAT 5721
1901 lIleGlnIleThrTyrValGluProTyrPheAspThrTyrGluMetLysAs 1917
|||||
5722 ATTCAATATACCTATGTCGAGCCATACCTTTGACACATATGATGAGGA 5771
1917 pArgIleThrTyrPheAspLysAsnTyrAsnLeuArgArgPheMetTyrC 1934
|||||
5772 CAGAAACACCTATTTCGACAAAATTACAACTCTGCTCATTCATGATCT 5821
1934 yThrProPheThrLeuAspGlyArgAlaHisGlyGluLeuHisGluGln 1950
|||||
5822 GTACACCCCTTACTTATAGATGGCCGTCAGGGGAACTTATGAAACA 5871
1951 PheLysArgLysThrIleLeuThrThrSerHisAlaPheProTyrIleLys 1967
5872 TTCAAAAAGGAAGCATCTCTGACTACGTCTCAATGCTTCCCTTAATTA 5921
1967 sThrArgValAsnValThrHisLysGluGluIleIleLeuThrProIleG 1984
|||||
5922 AACAAAGGTCATGTCACATTAAGAAAGATCATCTTAACACCAATTG 5971
1984 lValAlaIleGluAspMetGlnLysLysThrGlnGluLeuAlaPheAla 2000
|||||
5972 AAGTCTCTATGAGACATGACGAAAAAGACAGAGGATTTGACATTGGA 6021
2001 ThrHisGlnAspProAlaAspProLysMetLeuGlnMetValLeuGlnI 2017
6022 ACACATCAGGATCCGACAGACCCCAAAAGTCTTCAAGATGCTACTCAGG 6071
2017 ySerValGlyThrThrValAsnGlnGlyProLeuGlnValAlaGlnValP 2034
|||||
6072 ATCTGAGGCACACAGTGAATCAGGGGCTTTGGAAGTTGCCAGGTTT 6121
2034 heLeuSerGluIleProSerAspProLysLeuPheArgHisAsnLys 2050
|||||
6122 TTCTGTCTGAATATCCTAGTACCCAAAGCTCTTCAGACATCTAATTA 6171
2051 LeuArgLeuCysPheLysAspPheThrLysArgCysGluAspAlaLeuArg 2067
6172 CTGAGACTGTGCTTTAAAGATTTTACTAAAGGTGTGAAGATGCTTAAAG 6221
2067 gLysAsnLysSerLeuIleGlyProValGlnLysGlyTyrGlnArgGlu 2084
|||||

```

```

6222 AAAAATAGAGCTTAATTGGCCCGCTTCAAAAAGAGTATCAAAAGGAAT 6271
2084 euGIYtLysLeuSerSerPro 2090
|||||
6272 TGGGGAACATATCTTCGCCCT 6291

seq_name: /sids1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAH43851
seq_documentation_block:
ID AAH43851 standard; cDNA; 6372 BP.
XX
AC AAH43851;
XX
DT 04-SEP-2001 (first entry)
XX
DE Human CLASP-7 encoding cDNA sequence SEQ ID NO:1.
XX
KW Human; CLASP-7; cadherin-like asymmetry protein; immune response;
KW neuroprotective; antidiabetic; immunosuppressive; antirheumatic;
KW antiarthritic; hypotensive; anti-HIV; cytosstatic; immunostimulant;
KW antinaemic; antinflammatory; ophthalmological; nephrotrophic;
KW antihydroid; antiallergic; antiallergic; antiallergic; gene therapy;
KW chromosome 19q13.2; autoimmune disease; multiple sclerosis; toxemia;
KW juvenile diabetes; rheumatoid arthritis; pruritic urticarial papule;
KW hypertension; Rh incompatibility; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 13..6156
FT FT /*tag= a
FT FT /product= "CLASP-7"
XX
PN WO200142295-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-US34152.
XX
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
PR 11-APR-2000; 2000US-0196267.
PR 11-APR-2000; 2000US-0196460.
PR 11-APR-2000; 2000US-0196527.
PR 11-APR-2000; 2000US-0196528.
PR 11-APR-2000; 2000US-0547276.
PR 13-OCT-2000; 2000US-0240503.
PR 13-OCT-2000; 2000US-0240508.
PR 13-OCT-2000; 2000US-0240539.
PR 13-OCT-2000; 2000US-0240543.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu P, Garman JD, Candia AF;
XX
DR WPI: 2001-381641/40.
XX
DR P-PSDB; AAB99541.
XX
PT Novel cadherin-like asymmetry protein-7 and polynucleotides encoding
PT the polypeptide, useful for treating autoimmune disease,
PT hypersensitivity, preventing transplant rejection by modulating immune
PT response -
XX
PS Claim 1; Fig 5; 151pp. English.
XX
CC The present invention describes a human cadherin-like asymmetry protein
CC (CLASP), designated CLASP-7. The CLASP-7 protein (I) and its encoding
CC nucleotide sequence (II) have activities including: neuroprotective;
CC antidiabetic; immunosuppressive; antirheumatic; antiarthritic; anti-HIV;
CC hypotensive; cytosstatic; immunostimulant; antinaemic; antinflammatory;
CC ophthalmological; nephrotrophic; antihydroid; antiallergic; antiallergic;
CC antiallergic; and antibacterial. (II) and CLASP-7 antibodies (III) are

```

CC useful for detecting the CLASP-7 polypeptide. (ii) is useful for
CC producing (i) by recombinant methods. (i) or its fragment are useful for
CC inhibiting an immune response in a cell such as T cell or B cell. A
CC pharmaceutical composition (c), comprising (i) or (ii), can be useful
CC for treating CLASP-7-mediated disease such as an autoimmune disease
CC caused or exacerbated by increased activity of T helper cells. Autoimmune
CC diseases which can be treated using (c) include multiple sclerosis,
CC juvenile diabetes and rheumatoid arthritis. (i) is useful for treating
CC toxemia or pregnancy induced hypertension, pruritic urticarial papules
CC and Rh incompatibility. (i) is also useful as a diagnostic reagent for
CC immune and other disorders, since diseases characterised by
CC overproduction or depletion of lymphocytes in blood or other organs may
CC be detected by monitoring the level of (i) or its mRNA. CLASP-7 has been
CC mapped to the chromosomal location 19q13.2. The present sequence
CC encodes the human CLASP-7 protein as given in the present invention.
XX

SQ Sequence 6372 BP; 1273 A; 2050 C; 1849 G; 1200 T; 0 other;

alignment_scores:

Quality: 7136.50 Length: 2090
Ratio: 3.998 Gaps: 11
Percent Similarity: 85.407 Percent Identity: 65.311

alignment_block:

US-09-737-246-2 x AAH43851 ..

Align seg 1/1 to: AAH43851 from: 1 to: 6372

```
2 AlaGluArgArgAlaPheAlaGlnLysIleSerArgThrValAlaAlaG1 18
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
22 TCCGAGCCCGCGCTTCGCGACAGATCAACAGAGCGTGGCGCGAGA 71
uValArgpGlnIleSerGlyGlnTySerGlySerProGlnLeuLeuL 35
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
72 GGTGGGAAGCAGGTGTCCGGGAACGAGTGTCTCCCGACATCCAGCA 121
35 yAsnLeuAsnIleValGlnLysIleSerHisIsthrThrValProLeu 51
:::|||||:::|||||:::|||||:::|||||:::|||||
122 GG.....CGTCAGACAGCTCCCTGGGGGTCCCACTG 153
52 ThrGluAlaValAspProValAspLeuGlnAspTyIleuIlethrHispr 68
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
154 ACTGAAGTGTGAGCCCTGAGCTTGTGAGATGTACTTGTGAGCCGCC 203
68 OLeuAlaValAspSerGlyProLeuArgAspLeuIleGluPheProPro 85
|:::|||||:::|||||:::|||||:::|||||:::|||||
204 ACCAGATGCTAGAGCCCGCGCTCAGAGGACTGTGAAATTCACAGCTG 253
85 sPaspIleGluValTyIleSerProArgAspCysArgThrLeuValSer 101
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
254 ATGACTGTGAGCTGTGCTGAGCCCGGGAATGCGGACACAGGAGGCC 303
102 AlaValProGluGlnUserGluMetAspProHisValArgAspCysIle 118
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
304 GGGATCCCAAGGATGAAGAACTGATGCTGAGGCGCGGGGGA 353
118 gSerTyThrGluAspThrPalalIleValIleArgLysTyHisLysLeu 135
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
354 GATGTATATGTAGAGACTGGGTGATGTCCACAGAGTATCAGACTGA 403
135 LyrThrGlyPheAsnProAsnThrLeuAspLysGlnLysGluArgGlnLys 151
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
404 GTCAGCATACAGCCCGCTCAGCACACACAGCGGAGCGACAGAG 453
152 GlyLeuProLysGlnValPheGlnUserAspGluAlaProAspGlyLys 168
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
454 GGCCTCCCGCGCGCTTGTGAGAGATGCTTGTGAGAGCAGAGGTC 503
168 rTyGlnAspAspGlnAspAspLeuLysArgArgSerMetSerIleAsp 185
|:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
504 CGGCCCTGAGAGACTCGAATACCTCCGCGCTGGCTCCCGGAG 553
```

```
185 spthrProArgGlySerThrPalalCysSerIlePheAspLeuLysAsnSer 201
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
554 ACACCCCTGAAAGCAGTGTGCTTACATCTTCAGCTGAGAGACCTG 603
202 LeuProAspAlaLeuLeuProAsnIleLeuAspArgThrProAsnGln 218
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
604 GCAGCTGACATCTTGTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTG 653
218 uIleAspArgGlnAsnAspAspGlnArgLysSerAsnArgHisLysGlu 235
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
654 TGTGAGCCGCGCAATGAACCTTGCAGCGGAGCAGCCGCGCGCGCC 703
235 euPheAlaLeuHisProSerProAspGlnGluGluProIleGluArgLeu 251
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
704 TGCTACCCCTTACCGGCACTGAGAGATGAACCGCTGAGACGCTGT 753
252 SerValProAspIleProLysGlnHisPheGlnGluArgLeuValLys 268
|||:::|||||:::|||||:::|||||:::|||||:::|||||
754 AGCGGCCAGAGCCACCCGCGAGCACTTGGACAAAGATCTTGCTGCA 803
268 sCysLeuSerLeuLysPheGluIleGluIleGluProIlePheAlaSerL 285
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
804 GTGTGTGTGCTTCAGTGTGAGATGAATGAGCCCATCTTGGGATCT 853
285 euAlaLeuTyArgAspValLysGlnLysLysIleSerGlnAsnPheTy 301
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
854 TGCTCTGTATGATGTGCGGAGAAAGAGATCTCGAGAACTTGTAC 903
302 PheAspLeuAsnSerGlnGlnMetLysGlyLeuLeuArgProHisValPr 318
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
904 TTGACCTGATCGAGCTCGATCAAGAGGGGCTGCTCGGCGCTCAGG 953
318 oProAlaAlaIleThrThrLeuAlaArgSerAlaIlePheSerIleThr 335
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
954 CCACCTGCAATCTCCACCTGCGCGCTGCTCCATCTTCTGTGACT 1003
335 yProSerGlnAspValPheLeuValIleLysLeuGlnLysValLeuGln 351
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1004 ACCCTCAGCTGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1053
352 GlnGlyAspIleGlyLysLysLysLysLysLysLysLysLysLysLys 368
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1054 CAAGGAGATCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTG 1103
368 aAspAlaThrLysAsnLysGlnLysLysLysLysLysLysLysLysLys 385
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1104 GGACACAGCCAAAGAACAAAGAGACTAGAGAACTGCGCGCGCG 1153
385 sPdlPheCysGlnArgLeuGlyLysTyArgMetProPheAlaThrThr 401
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1154 AGCAGTGTGCAACCGCGCTGGCGCGCTACCGCATGCTTGGCTGGAG 1203
402 AlaIleHisLeuMetAsnIleValSerSerAlaGlySerLeuGluArg 418
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1204 GCCGTGCACTTGGCCACATCTGAGACAGCGCTGCGAGCTGAGCAG 1253
418 pSerThrGluValGluIleSerThrGlyLysArgLysGlySerThrP 435
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1254 CTCTGAC.....TGGAGGGCGAGCGCGCGCGAGCTGAGAG 1291
435 LuArgArgAsnSerSerIleValGlyArgArgSerLeuGluArgThrThr 451
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1292 ACCGGCGCGCT.....CGGGGCGCCAGAGCGGGCGAGT 1326
452 SerGlyAspAspAlaCysAsnLeuThrSerPheArgProAlaThrLeuTh 468
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1327 ACTGGGAGACAGCGCTGAGCTTCTGCTGCTGCTGCTGCTGCTG 1376
468 rValThrAsnPhePheLysGlnGluLysAspArgLeuSerAspGlnAsp 485
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1377 TGTCAAACTTCTTTAAGCAGAGAGCTGAGCGACTGATGACGAGGACC 1426
485 euTyLysPheLeuAlaAspMetArgArgProSerSerValLeuArgArg 501
```


[illegible]

1396	LeuGIuGIuAlaLeuLeuGlySerIleGlyAlaArgInGluMetValAr	1412	1397	CTGAGGAAACCCATTTCGTGTACCATCGGAGGTGCACAAAGATGGTCC	4022
1412	gATgSerArgGlyGluLeuGluArgSerProSerGlySerAlaPheGlyS	1429	1413	GCAGAGTCGT.....GAGGAGAGCCCG.....TTTGGGA	4051
1429	erGIuGluSLeuLeuArGTTPArgGlySAspMetThiHISrPArgGlnAsn	1445	1430	erGIuGluSLeuLeuArGTTPArgGlySAspMetThiHISrPArgGlnAsn	1445
4052	ATCCAGAGAACCGTCCCTGGCGGAAGAGGTGCACACACGTGAAAGCAAAAC	4101	4053	ATCCAGAGAACCGTCCCTGGCGGAAGAGGTGCACACACGTGAAAGCAAAAC	4101
1446	ThGCluGlySLeuAspLysSerArgAlaGluIleGluHISGluAlaLeuIle	1462	1447	ThGCluGlySLeuAspLysSerArgAlaGluIleGluHISGluAlaLeuIle	1462
4102	TCAGACCCCGCTGGACAAGACCAAGATGAAATGAGAACACGAGCCCTTGT	4151	4103	TCAGACCCCGCTGGACAAGACCAAGATGAAATGAGAACACGAGCCCTTGT	4151
1462	eAspGlyAsnLeuAlaThrGluAlaAsnLeuIleIleLeuAspThrLeuG	1479	1463	eAspGlyAsnLeuAlaThrGluAlaAsnLeuIleIleLeuAspThrLeuG	1479
4152	GGAGGAACACTGGCAACCGAGGAAAGCCTTAGTGTTGGACACACTGG	4201	4153	GGAGGAACACTGGCAACCGAGGAAAGCCTTAGTGTTGGACACACTGG	4201
1479	IuIleValAlaGlnPrrValSerValThrGluSerLysGluSerIleLeu	1495	1479	IuIleValAlaGlnPrrValSerValThrGluSerLysGluSerIleLeu	1495
4202	AGATCATCGGACGAGCGGTGATGCTTTCAAGAACCCGGGAGAGCGCTCTGG	4251	4202	AGATCATCGGACGAGCGGTGATGCTTTCAAGAACCCGGGAGAGCGCTCTGG	4251
1496	GIuGIYValLeuLysValLeuLeuHISSerMetAlaCysAsnGlnSerAl	1512	1496	GIuGIYValLeuLysValLeuLeuHISSerMetAlaCysAsnGlnSerAl	1512
4252	GGGCGAGTGTGAAGTTGTGCTGTACACCCGTGGCGACTGTCCACAGATGC	4301	4252	GGGCGAGTGTGAAGTTGTGCTGTACACCCGTGGCGACTGTCCACAGATGC	4301
1512	aValYrYrLeuGlnHIScysPheAlaThrGlnArgAlaLeuValSerLysP	1529	1512	aValYrYrLeuGlnHIScysPheAlaThrGlnArgAlaLeuValSerLysP	1529
4302	CCCTCTCTTGGCACACGAGCGTCCACCCAGAGGCGCTTGTCGTCACAGT	4351	4302	CCCTCTCTTGGCACACGAGCGTCCACCCAGAGGCGCTTGTCGTCACAGT	4351
1529	heProGluLeuLeuPheGluGluGluGluThrGluGlnCysAlaAspLeuGys	1545	1529	heProGluLeuLeuPheGluGluGluGluThrGluGlnCysAlaAspLeuGys	1545
4352	TCCCGAGCGCTGCTTCGAGAGAGACACGAGACTGTGTCCGACTGTGC	4401	4352	TCCCGAGCGCTGCTTCGAGAGAGACACGAGACTGTGTCCGACTGTGC	4401
1546	LeuArgLeuLeuArgHIScysSerSerIleGlyThrIleArgSerHIS	1562	1546	LeuArgLeuLeuArgHIScysSerSerIleGlyThrIleArgSerHIS	1562
4402	CTGAGGCTCTTACGACACGTGGACCGCGCATACACACTCCACGCA	4451	4402	CTGAGGCTCTTACGACACGTGGACCGCGCATACACACTCCACGCA	4451
1562	sProSerAlaSerLeuYrYrLeuLeuMetArgGlnAsnPheGluIleGlyA	1579	1562	sProSerAlaSerLeuYrYrLeuLeuMetArgGlnAsnPheGluIleGlyA	1579
4452	CGCGACGCGCTCGCTACCTCTCATCGACAGAACTTCAGATCGGC	4501	4452	CGCGACGCGCTCGCTACCTCTCATCGACAGAACTTCAGATCGGC	4501
1579	snaAsnPheAlaArgValLysMetGlnValProMetSerLeuSerSerLeu	1595	1579	snaAsnPheAlaArgValLysMetGlnValProMetSerLeuSerSerLeu	1595
4502	ACAACCTTCCCGGTGGAAGATGCAGAGTCACCATGTCTCTCGGCCCTG	4551	4502	ACAACCTTCCCGGTGGAAGATGCAGAGTCACCATGTCTCTCGGCCCTG	4551
1596	ValGIYThrSerGlnAsnPheAsnGluGluPheLeuArgArgSerLeuLys	1612	1596	ValGIYThrSerGlnAsnPheAsnGluGluPheLeuArgArgSerLeuLys	1612
4552	GTGGGAGACACGAGACTTCAGTAAAGACACCTGCACAGCTTCACTCAA	4601	4552	GTGGGAGACACGAGACTTCAGTAAAGACACCTGCACAGCTTCACTCAA	4601
1612	sThrIleLeuThrYrYrAlaGluAlaAspLeuValLeuArgGluThrThrP	1629	1612	sThrIleLeuThrYrYrAlaGluAlaAspLeuValLeuArgGluThrThrP	1629
4602	AACCATCTCACCTATGTGCGAGGAGACATGGGCGTGGGGACAGACACT	4651	4602	AACCATCTCACCTATGTGCGAGGAGACATGGGCGTGGGGACAGACACT	4651
1629	heProAspGlnValGlnAspLeuValPheAsnLeuHISMetIleLeuSer	1645	1629	heProAspGlnValGlnAspLeuValPheAsnLeuHISMetIleLeuSer	1645
4652	TCCCAAGACGAGGTCCAGGACCTCGATGTTCAACCTGCACATGATCTCTGACG	4701	4652	TCCCAAGACGAGGTCCAGGACCTCGATGTTCAACCTGCACATGATCTCTGACG	4701
1646	AspThrValLysMetLysGluHISGlnGlnLysAspProGluMetLeuIleAs	1662	1646	AspThrValLysMetLysGluHISGlnGlnLysAspProGluMetLeuIleAs	1662
4702	GACACGGTAAATGATGAGGACACACGAGAGACCTTGAGATGCTCATCGA	4751	4702	GACACGGTAAATGATGAGGACACACGAGAGACCTTGAGATGCTCATCGA	4751
1662	pleuMetYrArgIleAlaLysGlyTYrGlnThrSerProGlu...ArgL	1678	1662	pleuMetYrArgIleAlaLysGlyTYrGlnThrSerProGlu...ArgL	1678
4752	CCTCATGTCACAAATTCGCCGGGCTTACAGAGGCTCACCGACCTTGGCG	4801	4752	CCTCATGTCACAAATTCGCCGGGCTTACAGAGGCTCACCGACCTTGGCG	4801
1678	euThrTrpLeuGlnAsnMetAlaGlyLysHISSerGlnLysSerAspHis	1694	1678	euThrTrpLeuGlnAsnMetAlaGlyLysHISSerGlnLysSerAspHis	1694
4802	TGACCTGGTTGGAGAACATGGCGGGAGGACGCGGAGCTGGGCACAC	4851	4802	TGACCTGGTTGGAGAACATGGCGGGAGGACGCGGAGCTGGGCACAC	4851
1695	AlaGluAlaAlaGlnCysLeuValHISSerAlaAlaLeuValAlaGluTYr	1711	1695	AlaGluAlaAlaGlnCysLeuValHISSerAlaAlaLeuValAlaGluTYr	1711

```

|||||
4852 GCCGAGGCCGCCAGTCATGTCGACGCGCGCCCTCGTGCTGACGTA 4901
1711 rleuserMetLeuGluAspArgTyrLeuProValGlyCysValThrP 1728
|||||
4902 CTCGCCCTCTCGAGACCGACCGCCACTGCCCTGGCTGGCTTCCT 4951
1728 heGlnAsnIleSerSerAsnValLeuGluGluSerAlaValSerAsp 1744
|||||
4952 TCCAGAAACATCTCATCCACGTCGTAGAGAGTCCGCATCTCCAGCAG 5001
1745 ValValSerProAspGluGlyIleCysSerGlyLysTyrPheThrG 1761
|||||
5002 ATCTCTGCGCCGACGAGGAGGCTTCTCGCGGACGACCTTCACTGA 5051
1761 userGlyLeuValGlyLeuGluGluGlnAlaAlaSerPheSerMetA 1778
|||||
5052 GCTGGGGCTGGTAGGCTTGTGGACAGGAGCGGCTACTTCCATCGATG 5101
1778 laGlyMetTyrGluAlaValAsnGluValTyrLysValLeuIlePro 1794
|||||
5102 GCGGGCTCTACGAGCGGCGTAATGAGTCTACAGAACTCATCCCCATC 5151
1795 HisGluAlaAsnArgAspAlaLysLysLeuSerThrIleHisGlyLys 1811
|||||
5152 CTGGAAGCCGCCGCTACTACAGAAAGCTGCGCGGTGCACGGCAACT 5201
1811 uGlnGluAlaPheSerLysIleValHisGlnSerThrGlyTPGluArg 1828
|||||
5202 GCAGAGGCGCTTCACCAAGATCATGCACCAAGATTCGGCTGGAGCGG 5251
1828 ePheGlyThrTyrPheArgValGlyPheTyrGlyThrLysPheGlyAsp 1844
|||||
5252 TGTTCGGAGCATTTCCGCGTGGGCTTCTACGGCCGACCTTGGTGAC 5301
1845 LeuAspGluGlnGluPheValTyrLysGluProAlaIleThrLysLeu 1861
|||||
5302 CTGAGTACAGAGAGTTGTGTACAGAGACATCATGACCAAGAGCTGC 5351
1861 agIleSerHisArgLeuGluGlyPheTyrGlyGluArgPheGlyGlu 1878
|||||
5352 AGGATCTCACCGGCTGAGAGTTCTACAGGAGCAATTTGGCGAGC 5401
1878 sPValValGluValIleLysAspSerAsnProValAspLysCysLysLeu 1894
|||||
5402 AGTCGTGTGATATCAAAAGACTCTTACCTGTGACAAAGTCCAAAGCT 5451
1895 AspProAsnLysAlaTyrIleGlnIleThrTyrValGluProTyrPhe 1911
|||||
5452 GACTCACAAGGCGCTACATCCAGATCACGTATGTGAACCGTACTTGA 5501
1911 pThrTyrGluMetLysAspArgIleThrTyrPheAspLysAsnTyrAsn 1928
|||||
5502 TACCTACGAGCTCAAGACGGGTGACTTCTTGTGACCCCACTAAGGCG 5551
1928 euArgArgPheMetTyrCysThrProPheThrLeuAspGlyArgAlaHis 1944
|||||
5552 TTCGCAATTCCTGTCTGACGCGCTTCAAGCGGAGTGGCGCGACAC 5601
1945 GlyGluLeuHisGluGlnPheLysArgLysThrIleLeuThrThrSer 1961
|||||
5602 GGGGAGCTGCGCCGACACACAGCGTAAAGCGTCTCAGACCGACCA 5651
1961 saIaPheProTyrIleLysThrArgValAsnValThrHisLysGluGlu 1978
|||||
5652 GCGCTTCCCTACATCAAGACTCGCATCCGTGTGTCACCGGAGGAGA 5701
1978 leIleLeuThrProIleGluValAlaIleGluAspMetGlnLysLysThr 1994
|||||
5702 CGGTGCTGACGCGCATGAGAGTGGCATCGAGACATGACAAAGAGACA 5751
1995 GlnGluLeuAlaPheAlaThrHisGlnAspProAlaAspProLysMet 2011
|||||

```

```

5752 CGGAGCTGCGCTTGCACCGACGAGACCCACAGATCGTAAGATGCT 5801
2011 uGlnMetValLeuGlnGlySerValGlyThrThrValAsnGlnGlyPro 2028
|||||
5802 ACAAGTGTGCTTCAAGGCTCTGTAGGAGGAGGAGGAGGAGGAGGAG 5851
2028 euGluValAlaGlnValPheLeuSerGluIleProSerAspProLysLeu 2044
|||||
5852 TGGAGGTGGCCGAGGTGTTTGTAGACAGATCCGGAAGACCCAGCTC 5901
2045 PheArgHisHisAsnLysLeuArgLeuCysPheLysAspPheThrLys 2061
|||||
5902 TTCCGCGATCAACAACAAATGCGGCTCTGCTTCAAGGACTTCTCAAGA 5951
2061 GCysGluAspAlaLeuArgLysAsnLysSerLeuIleGlyProValGln 2078
|||||
5952 ATGTAGAGATGCGCTGCGGAAATAATAGGCCGTATGGCGGACAGCA 6001
2078 ysGluTyrGlnArgGluLeu 2084
|||||
6002 AGGATACACCGGTAGCTG 6021

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAH41953
seq_documentation_block:
ID AAH41953 standard; DNA; 4200 BP.
XX
AC AAH41953;
XX
DT 30-AUG-2001 (first entry)
XX
DE CLASP-3 nucleotide fragment.
XX
KW Human; cadherin-like asymmetry protein; CLASP; CLASP-3; immune response;
cell surface molecule; transmembrane protein; immunosuppressive; vaccine;
antitumor; anti-inflammation; antirheumatic; antidiabetic; dermatological; uropathic;
ophthalmological; nephrotropic; antithyroid; antidiabetic;
KW neuroprotective; antistimulant; antibacterial; antisense therapy;
KW gene therapy; chromosome 1p31.1; ds.
XX
OS Unidentified.
XX
PN WO200142297-A2.
XX
PD 14-JUN-2001.
XX
PE 13-DEC-2000; 2000WO-US34171.
XX
PF 13-DEC-1999; 90US-0170453.
XX
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
PR 11-APR-2000; 2000US-0196267.
PR 11-APR-2000; 2000US-0196460.
PR 11-APR-2000; 2000US-0196527.
PR 11-APR-2000; 2000US-0196528.
PR 11-APR-2000; 2000US-0547276.
PR 13-OCT-2000; 2000US-0240503.
PR 13-OCT-2000; 2000US-0240508.
XX
PA (ARBO-) ARBOR VITA CORP.
XX
PI Lu P, Garman JD, Candia AF.
XX
WPI; 2001-375003/39.
XX
DR Novel isolated cadherin-like asymmetry protein (CLASP)-3, useful for
PT useful in preventing or treating a CLASP-3-mediated disease e.g.
PT autoimmune disease
XX
PS Disclosure; Fig 9C; 189pp; English.
XX
CC The present invention describes an isolated polypeptide (I) comprising a
nucleotide (nt) sequence (Sia) that has at least 90% identity to the

```


1435 CAGATCTCTCACTTTGACAGCTAAACGGCTATTAGATCTGCTTATCTC 1484
1366 CysValSerCysPheGluTyrLysGlyLysValPheGluArgMetAs 1382
1485 TGTGTGCTTCTGCTTGAATATAAGGAAAAAGTCTTGAACGATATA 1534
1382 nSerLeuThrPheLysLysSerLysAspMetArgAlaLysLeuGluAla 1399
1535 TTAGCTTGACCTTTAAGAAATCAAAAGACATGAGCAAAAGCTTGAAGAAG 1584
1399 IaILeuleuGlySerIleGlyAlaArgGlnGluMetValArgArgSerArg 1415
1585 CTAATCTTGAGAGCATAGGTGCCAGGCAAGAAATGTAGCGGAGGAAC 1634
1416 GlyGlnLeuGluArgSerProSerGlySerAlaPheGlySerGlnGluAs 1432
1635 GGACAGCTCGAGAGAAAGCCATCTGGAAAGTGCCTTGGAAAGCAAGAAA 1684
1432 nLeuArgTyrPheGlyLysAspMetThrHisTyrArgGlnAsnThrGluLysL 1449
1685 TTTGAGGTGGAGAAAGATATGACTACCTGGCGTCMAAAACACAGAGAAC 1734
1449 euAspLysSerArgAlaGluIleGluHisGlyAlaLeuIleAspGlyAsn 1465
1735 TTGACAAATCAAGAGCAGACAGATTTGACACAGAGCAGTATGATGGAAC 1784
1466 LeuAlaThrGluAlaAsnLeuIleLeuAspThrLeuGluIleValAla 1482
1785 CTGGCTACAGAGCAACCTATATCATTTAGATACATTAAGATTTGTGT 1834
1482 IGLnThrValSerValThrGluSerLysGluSerIleLeuGlyGlyValL 1499
1835 TCAGACCTGTTCTGTACGGAATCCAAAGAGAGCATCTTGTGTGAGTGC 1884
1499 euLysValLeuLeuHisSerMetAlaCysAsnGlnSerAlaValTyrLeu 1515
1885 TMAAAGTCTGCTACTACACAGCATGGCTTAAACCAAGATGCAGTTATCTA 1934
1516 GlnHisCysPheAlaThrGlnArgAlaLeuValSerLysPheProGluLe 1532
1935 CAACACTGTTTGTCTACACAGAGCCTTGGTTCAAAAGTTCCGTGAACT 1984
1532 uLeuPheGluGluGluThrGluGlnCysAlaAspLeuCysLeuArgLeuL 1549
1985 CTTATTATAAGAGACAGAGACAGTGTGATTTATGCTCAGGCTTC 2034
1549 euArgHisCysSerSerSerIleGlyThrIleArgSerHisProSerAla 1565
2035 TCCGACACTGTAGCAGTAGCATCGGTACAATACGCTCACACCCACAGTCC 2084
1566 SerLeuTyrLeuLeuMetArgGlnAsnPheGluIleGlyAsnAsnPheAl 1582
2085 TCCCTTTACTACTAATAGGCCAAAACTTTGAGATTGGGAATTAACCTTGC 2134
1582 nArgValLysMetGlnValProMetSerLeuSerSerLeuValGlyThrS 1599
2135 CAGGGTTAAATGACGATACCAATGTCATCTCCCTTGGTGGGACACAT 2184
1599 eArgLAsnPheAsnGluGluPheLeuArgArgSerLeuLysThrIleLeu 1615
2185 CTCACAAATTTAATGACAAATCTTAAAGACGCTCTTAAACACTATATATG 2234
1616 ThrTyrAlaGluGluAspLeuGluLeuArgGluThrThrPheProAspG 1632
2235 ACATATGCTGAAGAAGATCTGAATTTAGGGAAACAAACATTTCCGTATCA 2284
1632 nValGlnAspLeuValPheAsnLeuHisMetIleLeuSerAspThrValL 1649
2285 GGTCCAGGATCTGGTTTCAATTCATATATGATCTTCTGTGATCTGTGA 2334
1649 yMetLysGluHisGlnGluAspProGluMetLeuIleAspLeuMetTyr 1665
2335 AAATGAAGAACACCAAGAGATCTCTGAATGTGATGATGATCTAATGTAC 2384

1666 ArgIleAlaLysGlyTyrGlnThrSerProGlu...ArgLeuThrTrpLe 1681
2385 AGAATTGCCAAGGCTTACAGACCTCTCCAGATCTGCGATTGACCTGGTT 2434
1681 uGlnAsnMetAlaGlyLysHisSerGluArgSerAsnHisAlaGluAla 1698
2435 GCAGAACATGCGACAGGCAAGCAGTACAGAACGACAAATCATCATCTGAAGCTG 2484
1698 IAGlnCysLeuValHisSerAlaAlaLeuValAlaGluTyrLeuSerMet 1714
2485 CACAGTGTCTAGTCCACTCAGACAGCACTTGTCTGTGAATATTTGAGCAATG 2534
1715 LeuGluAspArgLysTyrLeuProValGlyCysValThrPheGlnAsnI 1731
2535 CTGAGAGACCGGAATATCTTCTCGTGGATGTGTAAATTTACAGATAT 2584
1731 eSerSerAsnValLeuGluGluSerAlaValSerAspAspValValSerP 1748
2585 TTCATCTAATGTTTATAGAAATCTGCGGTCTCAGATGATGTGTTATCTC 2634
1748 rAspGluGluGlyIleCysSerGlyLysTyrPheThrGluSerGlyLeu 1764
2635 CAGATGAAGAAGTATCTGCTGGAAATATCTTACTGAGTCAAGACCTT 2684
1765 ValGlyLeuLeuGluGlnAlaAlaAlaSerPheSerMetAlaGlyMetTyr 1781
2685 GTGGGATTACTGACACAGACAGCTCTCTCTCATGTGCTGGCTGATGTA 2734
1781 rGluAlaValAsnGluValTyrLysValLeuIleProIleHisGluAla 1798
2735 TGAAGCAGTTAATGAAGTATACAAAGTATTTCTATTCATGAAAGCTA 2784
1798 sNArgAspAlaLysLysLeuSerThrIleHisGlyLysLeuGlnGluAla 1814
2785 ATCGGGATGCAAGAAACTATCCACAAATTCATGTAACCTTCAAGAGCA 2834
1815 PheSerLysIleValHisGlnSerThrGlyTyrGluArgMetPheGlyTh 1831
2835 TTGACGAAATTTGTATCATCAGATACGTGCTGGAGCGGATGTTGGCAC 2884
1831 rTyrPheArgValAlaGlyPheTyrGlyThrLysPheGlyAspLeuAspGluG 1848
2885 CTAATTTCTGCTGTTGTTTATGAAACCAAGTCCGGGATTTGATGAAC 2934
1848 IAGluPheValTyrLysGluProAlaIleThrLysLeuAlaGluIleSer 1864
2935 AAGATTTGTTACAGAGACCTGCAATTAACCAACTTGCACAGATATCT 2984
1865 HisArgLeuGluGlyPheTyrGlyGluArgPheGlyGluAspValAlaG 1881
2985 CACAGATTGAGAGGATTTACGAGAAAGATTTGAGAGAGATGTGGTTTA 3034
1881 uValIleLysAspSerAsnProValAspLysCysLysLeuAspProAsnL 1898
3035 AGTAATCAAAAGACTTAATCTCTAGACAAAGTGTAAATTAATGATCTCAACA 3084
1898 ySAlaTyrIleGlnIleThrTyrValGluProTyrPheAspThrTyrGlu 1914
3085 AGGCATATATTCAGATTACCTATGTGAGGCAATCTTGAACACATATGAG 3134
1915 MetLysAspArgIleThrTyrPheAspLysAsnTyrAsnLeuArgArgPh 1931
3135 ATGAGAGACAGAAATCACTATTTGACAAAAAATTAACAATCTTGTGATTT 3184
1931 eMetTyrCysThrProPheThrLeuAspGlyArgAlaHisGlyGluLeuH 1948
3185 CATGTACTGTACACCTTTACTTATGATGGCCGTGCCCATGGGGAAGCTTC 3234
1948 IsgGluGlnPheLysArgLysThrIleLeuThrThrSerHisAlaPhePro 1964
3235 ATGACAAATTCAAAAGAGAGACCAATTTCTGACTACGTCATGCTGCTTCT 3284

```

1965 TyrIleLysThrArgValAsnValThrHisLysGluGluIleLeuTh 1981
|||||
3285 TATATTAAACAGAGGTCAATGTCACTCATTAAGAGAGATCATCTTAAC 3334
|||||
1981 rProIleGluValAlaIleGluAspMetGlnLysLysThrGlnGluLeuA 1998
|||||
3335 ACCAATTGAGTGTCTATTGAGACATCAGAAAAAGACAGAGAGTTGG 3384
|||||
1998 lAphAlaIleThrHisGlnAspProIleAspProLysMetLeuGlnMetVal 2014
|||||
3385 CATTTGCACACATCAGATCCCGACACCCCAAAATGCTTCAGATGGTA 3434
|||||
2015 LeuGlnGlySerValGlyThrThrValAsnGlnGlyProLeuGluValAl 2031
|||||
3435 CTCACGGATCTGTGTGACACACAGTGAATCAGGGCCCTTTGGAAGTTC 3484
|||||
2031 aGlnValPheLeuSerGluIleProSerAspProLysLeuPheArgHisH 2048
|||||
3485 CCAGGTTTTCGTCTGAATACCTAGTACCCCAAGCTCTTCAGACATC 3534
|||||
2048 lAsnLysLeuArgLeuGlySerPheLysAspPheThrLysArgCysGluAsp 2064
|||||
3535 ATAAATAACTGCGACTCTGCTTTAAAGATTTTACTAAAAAGGTGGAAGAT 3584
|||||
2065 AlAlaLeuArgLysAsnLysSerLeuIleGlyProValGlnLysGluTyrG 2081
|||||
3585 GCCTTAAGAAAAAATAGACGCTATTATGGCCCGTTCAAAAAGAGTATCA 3634
|||||
2081 nArgGluLeuGlyLysLeuSerSerPro 2090
|||||
3635 AAGGGAATGGGGAAACTATCTTCGCCCT 3662
|||||
seq_name: /SIDSL/gcgdata/hold/-geneseq/geneseqn-emb1/NA2001A.DAT:AAH41911
seq_documentation_block:
ID AAH41911 standard; cDNA; 4143 BP.
XX
AC AAH41911;
XX
DT 30-AUG-2001 (first entry)
XX
DE Preliminary human CLASP-3 cDNA sequence Fig 1.
XX
KW Human; cadherin-like asymmetry protein; CLASP; CLASP-3; immune response;
KW cell surface molecule; transmembrane protein; immunosuppressive; vaccine;
KW antiinflammatory; antiarthritic; antianemic; dermatological; uropathic;
KW ophtalmological; antirheumatic; nephrotoxic; antidiabetic;
KW neuroprotective; antistimatic; antibacterial; antisense therapy;
KW gene therapy; chromosome 1p31.1; ss.
XX
OS Homo sapiens.
XX
PN W0200142297-A2.
XX
PD 14-JUN-2001.
XX
PE 13-DEC-2000; 2000WO-US34171.
XX
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
PR 11-APR-2000; 2000US-0196267.
PR 11-APR-2000; 2000US-0196460.
PR 11-APR-2000; 2000US-0196527.
PR 11-APR-2000; 2000US-0196528.
PR 11-APR-2000; 2000US-0547276.
PR 13-OCT-2000; 2000US-0240503.
PR 13-OCT-2000; 2000US-0240508.
XX
XX (ARBO-) ARBOR VITA CORP.
XX
XX Lu P, Garman JD, Candia AF;
XX

```

```

DR WPI, 2001-375003/39.
DR P-PSDB: AAB99493.
XX
PT Novel, isolated cadherin-like asymmetry protein (CLASP) -3, useful for
PT useful in preventing or treating a CLASP-3-mediated disease e.g.
PT autoimmune disease
XX
PS Example 5; Fig 1; 189pp; English.
XX
CC The present invention describes an isolated polypeptide (I) comprising a
CC nucleotide (nt) sequence (S1a) that has at least 90% identity to the
CC sequence given in AAH41934 and is immunologically cross-reactive with
CC the derived amino acid (aa) sequence (S1b) given in AAB99495 or shares a
CC biological function with native CLASP-3. (I) has immunosuppressive,
CC antiinflammatory, antiarthritic, antianemic, dermatological, uropathic,
CC ophtalmological, antirheumatic, nephrotoxic, antidiabetic,
CC neuroprotective, antiasthmatic and antibacterial activities, and can be
CC used in antisense therapy, vaccine production and gene therapy. CLASP-3
CC related sequences can be used in preventing or treating a CLASP-3
CC mediated disease, preferably an autoimmune disease by inhibiting an
CC immune response. The autoimmune disease is caused or exacerbated by
CC increased activity of T helper 1 lymphocytes (TH1). AAH41893 to AAH41953
CC and AAB99491 to AAB99507 represent sequences which are used in the
CC exemplification of the present invention. CLASP-3 is localised in the
CC chromosome location 1p31.1.
XX
SQ Sequence 4143 BP; 1251 A; 829 C; 890 G; 1173 T; 0 other;

alignment_scores:
Quality: 6157.50 Length: 1201
Ratio: 5.131 Gaps: 1
Percent Similarity: 99.917 Percent Identity: 99.750

alignment_block:
US-09-737-246-2 x AAH41911 ..
Align seg 1/1 to: AAH41911 from: 1 to: 4143

891 LeuAsnArgSerArgSerLeuSerAsnSerAsnProAspIleSerGlyTh 907
|||||
4 GTAAATGTTCTCGAAGCCTTAGTAATAGCAATCCAGATATATCTGGAC 53
|||||
907 rProThrSerProAspAspGluValArgSerIleIleGlySerLysGlyL 924
|||||
54 TCCACGCTCACAGATCATAGCTTCATCATCGGAGTAAGGTT 103
|||||
924 euAspArgSerAsnSerTrpValAsnThrGlyGlyProLysAlaAlaPro 940
|||||
104 TAGATCGCTCCAAATTCCTGGGTTAACACTGTGTCCAAAAGCTGCCCA 153
|||||
941 TrpGlySerAsnProSerProSerAlaGluSerThrGlnAlaMetAspR 957
|||||
154 TGGGATGCCAACCCAGTCCAGTGCAGATCAACACAGGCTATGGATGG 203
|||||
957 gSerCysAsnArgMetSerSerHisThrGluThrSerPheLeuGlnT 974
|||||
204 AAGTTGTAATCGTATGTCTTCGCACACAGAGAGCTCAAGTTCTTCANA 253
|||||
974 hTrLeuThrGlyAlaGluProThrLysLysLeuPheHisGluGluLeuAl 990
|||||
254 CATTAACGGGACGCTTACCAACTAAAAAGCTTTTTCACGAGGAGCTGGCT 303
|||||
991 LeuGlnTrpValValLysSerGlySerValArgGluSerAlaLeuGlnG 1007
|||||
304 TTGCAGTGGGTGTTTGCAGTGGCAGGCTGGGAAATCAGCTTTGCAACA 353
|||||
1007 nAlaTrpPhePhePheGluLeuMetValLysSerMetValHisHisLeuT 1024
|||||
354 AGCGTGTTCTTTTTCATTAATGTAAGACATGTCACACATTAT 403
|||||
1024 yrPheAsnAspLysLeuGluAlaProArgLysSerArgPheProGluArg 1040
|||||

```


404 ACTTAATGATTAACCTTGAGGCTCCAGAGAAAAGTGTTCACAGACGT 453
1041 PheMetAspAspIleAlaIleuValSerThrIleAlaSerAspIleVal 1057
|||||
454 TTCAATGATGATGACATTCGACCTTTCTGACGACGATTGCTAGTGAATAGT 503
1057 lSerArgPheGlnLysAspThrGluMetValGluArgLeuAsnThrSerL 1074
|||||
504 TTCACGATTTTCAGAGAGACACAGAAATGTTGAGAGACTCAATACAAACC 553
1074 euAlaPhePheLeuAsnAspLeuLeuSerValMetAspArgGlyPheVal 1090
|||||
554 TTGCATCTCTTCTCAATGATCTGTGTCTGTATGACACAGAGATTTGTT 603
1091 PheSerLeuIleLysSerCysTyrLysGlnValSerSerLysLeuTyrSe 1107
|||||
604 TTTAGCCTTATAAAGTCTCGCTATTAACAGGTGCTTCAAGCTTTACTC 653
1107 rLeuProAsnProSerValIleuValSerLeuArgLeuAspPheLeuArgI 1124
654 ATTACCGAATCCCAAGTGTCTGTGCTGCTGAGGCTGGATTTTTCAGAA 703
1124 lIleLysSerHisGlnHisTyrValThrLeuAsnLeuProCysSerLeu 1140
|||||
704 TCATCTGCAGTCATAGCACATATGTATTAACCTTACCCTGCAGCTTA 753
1141 LeuThrProAlaSerProSerProSerValSerSerAlaThrSerGly 1157
|||||
754 CTTCCTCCACCTGCATCTCCATCCATCCCTCTGTCTCTGCAACATCTCA 803
1157 nSerSerGlyPheSerThrAsnValGlnAspGlnLysIleAlaAsnMetP 1174
804 GAGTCTCGATTTTTCAGCAATGTACAAGACCAGCAAAAGATTGCAGAAATATGT 853
1174 heGluLeuSerValProPheArgGlnGlnHisTyrLeuAlaGlyLeuVal 1190
|||||
854 TTGAATTAATCCGTGCTTCCTCCCAACACACATATTTGGCAGACTTGTG 903
1191 LeuThrGluLeuAlaValIleLeuAspProAspAlaGluGlyLeuPheGly 1207
|||||
904 TTAAAGAGACTGGCTGTCTCATTTTAAAGCCCTGATGCTGAAGAGCTGTGG 953
1207 yLeuHisLysLysValIleAsnMetValHisAsnLeuLeuSerSerHisA 1224
954 ATTGCATTAAGAAAGTCATCAATATGTTACACAATTACTCTCCAGTCAAG 1003
1224 spSerSerProArgTyrSerAspProGlnIleLysAlaArgValAlaMet 1240
1004 ACTCAGACCCGCGGTACTCTGACCTCAGATTAAGGCTCGAGTGGCCAG 1053
1241 LeuTyrLeuProLeuIleGlyIleIleMetGluThrValProGlnLeuTyr 1257
1054 TTGTATCTATCTCTGATTTGATTTATCAAGAAACTGTACTCAGCTGTA 1103
1257 rAspPheThrGluThrHisAsnGlnArgGlyArgProIleCysIleAlaT 1274
1104 TGATTTTACAGAAACTCAATCAACAGAGAGAACCAATTTGTATAGCCA 1153
1274 hrAspAspTyrGluSerGluSerGlySerMetIleSerGlnThrValAla 1290
1154 CTGATGATTAATGAAGTAGAGCGGAGAGATATGATAGCCAGACCGTTGCC 1203
1291 MetAlaIleAlaGlyThrSerValProGlnLeuThrArgProGlySerPh 1307
1204 ATGGCAATGCGAAGGACATCGGTCCCTCAACTAACAAGGCTGGCAGTTT 1253
1307 eLeuLeuThrSerThrSerGlyArgGlnHisThrThrPheSerAlaGlnLys 1324
1254 CCCTCTCAGGTCAACGAGTGGCAGCAACACACTACTTTCAGCAGAAAT 1303
1324 eTsrerArgSerLeuLeuIleCysLeuLeuTyrValLeuLysAsnAlaAsp 1340
1304 CAAAGTGAAGCCTTTTGATCTGTCTACTTTGGGTTCTCAAAAATGCAAGAT 1353

1341 GluThrValLeuGlnLysTrpPheThrAspLeuSerValLeuGlnLeuAs 1357
1354 GAACACATTTCTACAGAAAGTGTTTACAGATCTCTCAGCTGTGACGCTAA 1403
1357 nArgLeuLeuAspLeuLeuTyrLeuCysValSerCysPheGluTyrLysG 1374
1404 CCGGCTATTGATCTGCTTATCTCTGTGTGTCTGTCTGTGATATTAAG 1453
1374 lYlyLysValPheGluArgMetAsnSerLeuThrPheLysLysSerLys 1390
1454 GGAATAAAGTGTGGAAGATGAATGATGCTTGACCTTTAAGAAATCAAAA 1503
1391 AspMetArgAlaLysLeuGlnGluAlaIleLeuGlySerIleGlyAlaAr 1407
1504 GACATGAGACCAAAAGCTTGAAGAAAGCTATTCTTGGAGACGTAGTGCCAG 1553
1407 gGlnGluMetValArgArgSerArgGlyGlnLeuGluArgSerProSerG 1424
1554 GCAAGAAATGATGATCGCGGAGCCGAGGACAGCTCGAGAGAACCCATCTG 1603
1424 lYSerAlaPheGlySerGlnGluAsnLeuArgTyrPargLysAspMetThr 1440
1604 GAACTGGCTTTGGAAATCAAGAAATTTGAGGTGAGAGAAAGATATGACT 1653
1441 HisTrpArgGlnAsnThrGluLysLeuAspLysSerArgAlaGluIleGly 1457
1654 CACTGGGCTCAAAACACAGAGAACCTTGACAAATCAAGAGACAGATTTGA 1703
1457 uHisGluAlaLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeuIleT 1474
1704 ACACAGAGACACTGATGATGAAACCTGCGCTACAGAGCAAACTATATCA 1753
1474 lLeuAspThrLeuGluIleValValGlnThrValSerValThrGlnGlnSer 1490
1754 TTTTGAATACATTAGAGATTTGTTGTTGACACCGTTTCTGTAACGGAATTC 1803
1491 LysGlnSerIleLeuGlyGlyValLeuLysValLeuLeuHisSerMetAl 1507
1804 AAAGAGAGCATTTCTGGTGGAGTGTCAAAAGTGTACTATACAGACATGGC 1853
1507 aCysAsnGlnSerAlaValTyrLeuGlnHisCysPheAlaThrGlnArgA 1524
1854 CTGTAAACCAAGTGCAGTTTATCTACAAACACTGTTTGTCTACACAGAGAG 1903
1524 lAlaValSerLysPheProGluLeuLeuPheGluGluGluThrGlnGln 1540
1904 CTTTGGTTTCAAAAGTTTCTGAACTCTTATTGTAAGAGAGACAGACAG 1953
1541 CysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSerIleGly 1557
1954 TGTGCTGATTTATGCCACAGGCTTCTCCGACACTGTAGCAGTGAACGCG 2003
1557 yThrIleArgSerHisProSerAlaSerLeuTyrLeuLeuMetArgGlnA 1574
2004 TACAAATACGGTCAACACCCAGTGCCTCCCTTACCTACTAATGAGGCCAA 2053
1574 snPheGluIleGlyAsnAsnPheAlaArgValLysMetGlnValProMet 1590
2054 ACTTTGACATTTGGAAATTAATCTTCCAGGCTTAAATGAAGCAGTACCAATG 2103
1591 SerLeuSerSerLeuValGlyThrSerGlnAsnPheAsnGlnGluPheLe 1607
2104 TCACATATCTCTCTGTGGGACATCTCAGAAATTTTAAAGAATATCTT 2153
1607 uArgArgSerLeuLysThrIleLeuThrTyrAlaGluGlnAspLeuGluL 1624
2154 AAGAGCTTCTCTTAAAGACTATATGACATATGCTGAAGAGATCTGGAAT 2203
1624 euArgGluThrThrPheProAspGlnValGlnAspLeuValPheAsnLeu 1640
2204 TGAAGGAACACAAATTTCTGATCAGGTCCAGAGATCTGTGTTCAATCTC 2253

```

1641 HISMELLEuSerAspThrValIysMetLysGIuHISGInGIuAspPr 1657
|||||
2254 CATATGATCTTCTGATCTGTAAGAAAGAGAACACAGAGAGATCC 2303
|||||
1657 OGluMetLeuIleAspLeuMetTyrArgIleAlaLysGlyTyrGlnThrS 1674
|||||
2304 TGAAATGTTGATGATCTAATGATACAGAAATGCGAAGGGTTACGAGACT 2353
|||||
1674 ePProGlu...ArgLeuThrTrpLeuGInAsnMetAlaGlyLysHisSer 1689
|||||
2354 CTCCAGATGCGGATGACCTGGTGGCAGAACATGGCAGGCAAGCATCA 2403
|||||
1690 GluArgSerAsnHisAlaGluIleAlaGInCysLeuValHisSerAlaI 1706
|||||
2404 GAACGAGCAATCATGCTGAACCTGCACAGTGTAGTCCACTCGACGAC 2453
|||||
1706 AleuValAlaGluTyrLeuSerMetLeuGluAspArgLysTyrIleuPro 1723
|||||
2454 ACTTGTGCTGAATATTGAGCATGCTGGAGACCGGAAATATCTTCTG 2503
|||||
1723 aGICysValThrPheGlnAsnIleSerSerAsnValLeuGInGluSer 1739
|||||
2504 TGGGATGCTTAACATTCAGAAATATTCATCTAATGTTTGAAGAAATCT 2553
|||||
1740 AlaValSerAspAspValValSerProAspGluGluGlyIleCysSerG 1756
|||||
2554 GCGGCTCAGATGATGGTATCTCCAGATGAAGAAAGATCTGCTGG 2603
|||||
1756 yLysTyrPheThrGluSerGlyLeuValGlyLeuLeuGluGlnAlaIle 1773
|||||
2604 AAAAATCTTACTGAGTCAGGACTTGGGATTTACTGGAACACAGCAGCTG 2653
|||||
1773 lAsSerPheSerMetAlaGlyMetTyrGluAlaValAsnGluValTyrLys 1789
|||||
2654 CTTCCTCTCTGCTGGCTGGCATGTAAGAAGCATTAATGAAGTTTACAAA 2703
|||||
1790 ValLeuIleProIleHisGluAlaAsnArgAspAlaLysLysLeuSerTh 1806
|||||
2704 GTACTTATTCCTATTCATGAACCTAATCGGATGCAAGAAACCTATCCAC 2753
|||||
1806 rIleHisGlyLysLeuGInGluAlaPheSerLysIleValHisGlnSerT 1823
|||||
2754 AATTCATGTAACCTTCAGAGCAATTCAGCAAAATGTTATCATCGAGCTA 2803
|||||
1823 hrcGlyTrpGluArgMetPheGlyThrTyrPheArgValGlyPheTyrGly 1839
|||||
2804 CTGGCTGGAGCGGATGTTGGCACCCTATTTCCGTGGTTTATATGCA 2853
|||||
1840 ThrLysPheGlyAspLeuAspGluGInLupheValTyrLysGluProAl 1856
|||||
2854 ACCAGATTTCGGGATTTGGATGACAAACAATTTGTTTACAAGAGCCCTGC 2903
|||||
1856 alleThrLysLeuAlaGluIleSerHisArgLeuGInGlyLysGlyArg 1873
|||||
2904 AATTAACCAACTTCAGAGATATCTCACAAGATGGAGGGATTTTACGGAG 2953
|||||
1873 LuArgPheGlyGluAspValAlaGluValIleLysAspSerAsnProVal 1889
|||||
2954 AAAGATTTGGAGAGCATGTGGTTGAAGTAATCAAAAGACTCAATCCCTCTA 3003
|||||
1890 AspLysCysLysLeuAspProAsnLysAlaTyrIleGlnIleThrTyrVa 1906
|||||
3004 GACAGATGTAAATATGATCTTAACAAGCATATATATTCAGATTACCTATGT 3053
|||||
1906 IGluProTyrPheAspThrTyrGluMetLysAspArgIleThrTyrPheA 1923
|||||
3054 GGAGCCATACCTTGACACATATGAGATGAAGACAGAACTACCTATTTGCG 3103
|||||
1923 sPLysAsnTyrAsnLeuArgArgPheMetTyrCysThrProPheThrLeu 1939
|||||
3104 ACAAATTTACATCTTGTGTCGATTCATGTAAGACCCCTTACTTTA 3153
|||||
1940 AspGlyArgAlaHisGlyLeuLeuHisGluGlnPheLysArgLysThrII 1956

```

```

|||||
3154 GATGGCCGTCGCCCATGGGGAATCTCATGAACAATTCAAAAGGAAGCAT 3203
|||||
1956 eLeuThrThrSerHisAlaPheProTyrIleLysThrArgValAsnValT 1973
|||||
3204 TCTGACTACGCTCATGCTCTTCCTTATATTAAACAAAGGGTCAATGTCA 3253
|||||
1973 hrcHisLysGluGluIleIleLeuThrProIleGluValAlaIleGluAsp 1989
|||||
3254 CTCTTAAGAAGAGATCATCTTAAACACCAATGTGAAGTGTCTATTGAGAGC 3303
|||||
1990 MetGlnLysLysThrGInGluLeuAlaPheAlaThrHisGlnAspProAl 2006
|||||
3304 ATGAGAGAAAAGACAGACAGAGTTGGCATTTGCCAACATCAGATCCCGC 3353
|||||
2006 aAspProLysMetLeuGInMetValLeuGInGlySerValGlyThrThy 2023
|||||
3354 AGACCCCAAAATGCTTCAGATGGTACTCCAGGATCTGTAGCCACACAG 3403
|||||
2023 aLAsnGInGlyProLeuGluValAlaGInValPheLeuSerGluIlePro 2039
|||||
3404 TGAATCAGGGGCCCTTTGGAAAGTTGCCAGGTTTCTGTCTGAATATCCT 3453
|||||
2040 SerAspProLysLeuPheArgHisHisAsnLysLeuArgLeuCysPheL 2056
|||||
3454 AGTGACCCCAAGGCTCTTCACACATCATTAATAACTCGACTCGCTTTAA 3503
|||||
2056 sAspPheThrLysArgCysGluAspAlaLeuArgLysAsnLysSerLeuI 2073
|||||
3504 AGATTTTACTTAAAGAGGTGAAGATGCTTAAAGAAAAATTAAGAGCTTAA 3553
|||||
2073 lGluProValGlnLysGluTyrGlnArgGluLeuGluLysLeuSerSer 2089
|||||
3554 TTGGGCGCGGTTCAAAAGAGATATCAAAAGGAATTTGGGAAACTATCTTGG 3603
|||||
2090 Pro 2090
|||
3604 CCT 3606

seq_name: /STD1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAH41912
seq_documentation_block:
ID AAH41912 standard; cDNA; 4144 BP.
AC AAH41912;
XX
DT 30-AUG-2001 (first entry)
DE
XX Human CLASP-3 cDNA sequence Fig 4.
XX
DE Human: cadherin-like asymmetry protein; CLASP; CLASP-3; immune response;
KW cell surface molecule; transmembrane protein; immunosuppressive; vaccine;
KW antiinflammatory; antiarthritic; antiinfective; dermatological; uropathic;
KW ophthalmological; antirheumatic; nephrotoxic; antidiabetic;
KW neuroprotective; antiaslatic; antibacterial; antisense therapy;
KW gene therapy; chromosome 1p31.1; ss.
XX
OS Homo sapiens.
XX
PN WO200142297-A2.
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-US34171.
XX
PR 13-DEC-1999; 99US-0170453.
PR 14-JAN-2000; 2000US-0176195.
PR 14-FEB-2000; 2000US-0182296.
PR 11-APR-2000; 2000US-0196267.
PR 11-APR-2000; 2000US-0196460.
PR 11-APR-2000; 2000US-0196527.
PR 11-APR-2000; 2000US-0196528.
PR 11-APR-2000; 2000US-0547276.

```

13-OCT-2000; 2000US-0240503.
13-OCT-2000; 2000US-0240508.
(ARBO-) ARBOR VITA CORP.
Lu P, Garman JD, Candia AF;
WPI: 2001-375003/39.
P-PsDB: AAB99494.
Novel isolated cadherin-like asymmetry protein (CLASP)-3, useful for
autoimmune disease -
Disclosure; Fig 4A; 189pp; English.
The present invention describes an isolated polypeptide (I) comprising a
nucleotide (nt) sequence (SIa) that has at least 90% identity to the
sequence given in AAH41934 and is immunologically cross-reactive with
the derived amino acid (aa) sequence (SIb) given in AAB99495 or shares a
biological function with native CLASP-3. (I) has immunosuppressive,
antiinflammatory, antiarthritic, antianaemic, dermatological, uropathic,
neuroprotective, antihemetic, nephrotropic, antihypoid, antidiabetic
used in antisense therapy, vaccine production and gene therapy. CLASP-3
related sequences can be used in preventing or treating a CLASP-3
mediated disease, preferably an autoimmune disease by inhibiting an
immune response. The autoimmune disease is caused or exacerbated by
increased activity of T helper 1 lymphocytes (TH1). AAH41934 to AAH41953
and AAB99491 to AAB99507 represent sequences which are used in the
exemplification of the present invention. CLASP-3 is localised in the
chromosome location 1p31.1.

[illegible]

```
alignment_block:
US-09-737-246-2 x AAH41912 ..
```

Align seg 1/1 to: AAH41912 from: 1 to: 4144

891 LeuAsnArgSerArgSerLeuSerAsnSerAsnProAspIleSerGlyThr 907
 :::
 5 GTAATCGTCTCGAGCGCTAGTATATGCAATCCAGATATATCTGGAC 54

907 rprothrserproaspsgluvalargserileileglyserlysglyl 924
 |||||
 55 tccacagtcaccagatgatgaagttcgatcaatcatcggagtaagggtt 104

924 euaspargSerAsnSerTpvalasnthrglyglyProlysalaalapro 940
 |||||
 105 TAGATCGCTCCATTCCTGGGTTAACACTGGTGGTCCAAAGCTGCCCCA 154

941 TTPGlySerAsnProSerProSerAlaGluSerThrGlnAlaMetAspArg 957
 |||||
 155 TGGGATCCCAACCCAGTTCAGATCAACACAGGCTATGATCG 204

957 gSerCysasnArgmetSerSerHisThrGluThrSerSerPheLeuGlnT 974
 |||||
 205 AACGTGTATATCGTATGCTTCGCACACAGAGAGACGTCACAGTTCTTACAA 254

974 hrlcuthrglyargleuprothrlysllysleuphehisglululeuala 990
|||||
255 catpaaaggagagcttaccaactaaaaagctttttcacgagagagctggct 304

```

991 LeuGIInTPValValCysSerGlySerValArgGIuSerAlaLeuGIInGI 1007
|||||
305 TTGCAGTGGGTTGTTGCAGTGGCAGCGTTCCGGGATCAGCTTTGCACAA 354

```

1007 malatrphephegluleumetvallyssermetvalhishisleut 1024
 |||||
 355 agccgtctctttttgatattatgtaagagacatggtgcaccatttat 404

1024 yrphasnaspLysLeuGluaIaProaRgLySseraRgPheProGluArg 1040
405 ACTTTATGATAAACTTGAGGCTCCAAGAAAAAGTCGTTTCCAGAACGT 454

1041 phemetaspsp1lealaaleuvalserThrlealaSerasp1leua 1057
 |||||
 455 TTCATGATGACATTGCAGCTCTTGTACACAGATTGCTAGTATAGT 504

1057 IserArgpHeGInLysAspThrGluMetValGluArgLeuAsnThrSerL 1074
|||||
505 TTCACGATTTTCAGAGGACACAGAAATGCTTGAGAGACTCAATACAGCC 554

1074 euAlaPhePheLeuAsnAspLeuSerValMetaspArgGlyPheVal 1090
|||||
555 TTGCATCTTTCTCATGATGATCTGTCTCTATTATGACAGAGATTTCGT 604

1091 PheserleuilelyssercystylrlysginValserSerlyslleuTyrsE 1107
 |||||
 605 TTTAGCCTTATAAGTCCTGCTATAACACAGTGCTTCAAGCTTTACTC 654

```

1107 rleuprAsnProSerValLeuValSerLeuArgLeuAspPheLeuArgI 1124
      |||
655 ATTACCGATCCAGTGTTCTGGTGTCCCTTGAGGCTGCATTTCACGAA 704

```

```

1124 lelleCysserHisGluHistyValThrLeuAnsLeuProCysSerLeu 1146
      |||||
705 TCATCTGCAGTCATGAGCACTATGTACATTTAACTTACCCTGCAGCTTA 754

```

1141 LeutnPrProAlaserProserProSerValserSeralathrSergl 1157
|||||
755 CTTACTCCACCTGCATCTCCATCACCCTTCGTCTCTCGAACAATCTCA 804

```

1157 nSerSerGlyPheSerThrAsnValGlnAspGlnLysIleAlaAsnMetP 1174
      |||||
      |||||
      |||||
805 GAGTTCGATTTTCTACGATGTACACAGACCAAGAATTGCAATATATGT 854

```

```

1174 hegluleUserValPropheargInghIstYrleuAlaglyLeuVal 1190
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
855 TTGATTTATTCGTGCCTTCCGCCAACAGCATTAATTGGCAGACTTG TG 904

```

1191 LeuTrnGluLeuAlaValIleLeuAspProAspAlaGluGlyLeuPheG1 1207
 |||||
 905 TTACAGACGCTGCTGCATTTTAGACCGCTGATGCTGAAGACTGTTTG 954

```

1207 yLeuHISLysLysValILeasmetValHisasnLeuLeuSerSerHisA 122
      |||||
955 ATTGCATAGCAAGTCATCATATATGGTACACAACTTTACTCTCCAGTCACG 100

```

1224 spSerAspProArgTyrSerAspProGlnIleLysAlaArgValAlaMet 1244
 |||||
 1005 ACTCAGACCCCGCGTACTCTGACCCCTCAGATTAAGGCTCGAGTGGCCATG 105

1241 LeuTyrLeuProLeuIleGlyIleIleMetGluThrValProGlnLeuTy 125
 |||||
 1055 TTGTATCTACCTCTGATGTGATATATCATGGAACCTGTACCTCAGCTGTA 110

1257 rAspPheThrGluThrHisAsnGlnArgGlyArgProIleCysIleAlaT 127
1105 TGATTTTACAGAACTCACATTCACGAGAGACACCAATTGTGTATAGCCA 115

```

1274 hraspaspytyrglusergluserglysermetileserglnthrValAla 1299
      |||||
1155 CTGATGATTATGAAGTGAGAGCGGAAGTATGATTAAGCCAGACCGTTGCC 1200
      |||||

```

```

1291 MetaAilealgiYthrSeraProglInleutharProglySerPh 1307
      |||
1205 ATGCATCGCAGGACATCGGTCCCTCACTAACAGGCTGCGAGTTT 1255

```

```
1307 eLeuLeuThrSerThrSerg1yArgGlnHisThrThrPheSerAlaGln 1324
|||||
1255 CCTCTCAGCTCAACGAGTGGCAGCAACACTACTTTCAGCAGAT 1304
|||||
1324 eTerSerSerLeuLeuIleCysLeuLeuTrpValLeuLysAsn1aasp 1340
|||||
1305 CAAGTGAAGCCCTTTGATCTGCTACTTGGGTTCTCAAAAATGAGAT 1354
|||||
1341 GluThrValLeuGlnLysTrpPheThrAspLeuSerValLeuGlnLeu 1357
|||||
1355 GAACAGCTTCTACAGAGTGCTTACAGATCTCTCAGTCTTGACGTAA 1404
|||||
1357 nArgLeuLeuAspLeuLeuTyLeuCysValSergCysPheGluTrpLysG 1374
|||||
1405 CCGGCTATTAGATCTGCTTATCTCTGTGTCTTCTTGAGTAAAG 1454
|||||
1374 LysLysLysValPheGluArgMetAsnSerLeuThrPheLysLysSerLys 1390
|||||
1455 GGAATAAGCTTTGAGCAATGAATGACTTGACTTAAAGAAATCAAAA 1504
|||||
1391 AspMetArgAlaLysLeuGlnGluAlaIleLeuGlySerIleGlyAlaAr 1407
|||||
1505 GACATGAGCAAAAGCTTGAAGAGCTATCTTGGAGCATAGTGCAG 1554
|||||
1407 gGlnGluMetValArgArgSerArgGlyGlnLeuGluArgSerProSerg 1424
|||||
1555 GCAGAATAATGTATCGCGAAGCCGAGACAGCTCGAGAGAACCCATCG 1604
|||||
1424 LysSerLaaPheGlySerGlnGluAsnLeuArgTrpArgLysAspMetThr 1440
|||||
1605 GAAGTCCCTTGGAACTCAAGAAATTTGAGGTGGAGAAAGATATGACT 1654
|||||
1441 HistPrArgGlnAsnThrGluLysLeuAspLysSerArgAlaGluIleG 1457
|||||
1655 CACTGCGCTCAAAACACAGAGAAAGCTTGCAAAATCAAGAGCAGAAATTA 1704
|||||
1457 uHisGluAlaLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeuIle 1474
|||||
1705 ACACGAGACACTGATGTATGATGAACCTGCTACAGAACCAACTATATCA 1754
|||||
1474 IeLeuAspThrLeuGlnIleValGlnThrValSerValThrGluSer 1490
|||||
1755 TTTTAAATACATTAAGATTTGTTGACAGCCGTTTCTGTAAAGGAATCC 1804
|||||
1491 LysGluSerIleLeuGlyGlyValLeuLysValLeuLeuHisSermetal 1507
|||||
1805 AAAGAGAGATTTCTGTGAGACTGCTAAAGTCTACTACACAGCATGCG 1854
|||||
1507 acYsAsnGlnSerAlaValTyLeuGlnHisCysPheAlaThrGlnArgA 1524
|||||
1855 CTGTAACTCAAGTGCAGTTTATCTAACACACTGTTTGTCTACACAGAG 1904
|||||
1524 IalaValSerLysPheProGluLeuLeuPheGluGluThrGlnGln 1540
|||||
1905 CCTTGTTTCAAAAGTTTCTGAAGCTTATTTTGAAGAAGACAGAGAG 1954
|||||
1541 CysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSerIleG 1557
|||||
1955 TGTGCGATTTATGCTCAGGCTTCTCCGACACTACAGATGATGCG 2004
|||||
1557 yThrIleArgSerHisProSerAlaSerLeuTyLeuLeuMetArgGlnA 1574
|||||
2005 TACAAATACGGTACACACCCAGTGCCTCCCTTACTACTAATGAGGCA 2054
|||||
1574 snPheGlnIleGlyAsnAsnPheAlaArgValLysMetGlnValProMet 1590
|||||
2055 ACTTTGAGATTGGAAATACTTTGCCAGGGTTAAATGACAGTACCAATG 2104
|||||
1591 SerLeuSerSerLeuValGlyThrSergLysAsnPheAsnGluGluPhe 1607
|||||
2105 TCACATCTCTCTGTGTGGGCACATCTCAGAATTTTATGAAGAATTTCTT 2154
|||||
1607 uArgArgSerLeuLysThrIleLeuThrTyAlaGluGluAspLeuGln 1624
|||||
|||||
2155 AAGAGTCTCTAAAGATATATGACATATGCTGGAAGAGATGTGCAAT 2204
|||||
1624 euArgGluThrThrPheProAspGlnValGlnAspLeuValPheAsnLeu 1640
|||||
2205 TGAGGAAACAACATTTCTGATCAGGTCCAGATCTGTGTTTCAATCTC 2254
|||||
1641 HisMetIleLeuSerAspThrValLysMetLysGlnHisGlnGluAspPr 1657
|||||
2255 CATATGATTTCTTCTGATCTGTAATAATGAGGAAACACAGAGGATGCC 2304
|||||
1657 oGluMetLeuIleAspLeuMetTyArgIleAlaLysGlyTyGlnThrS 1674
|||||
2305 TGAATGTGTATGATCATTAATGTACAGAAATGGCAAGGGTTACAGACCT 2354
|||||
1674 erProGlu...ArgLeuThrTrpLeuGlnAsnMetAlaGlyLysHisSer 1689
|||||
2355 CTCAGATCTGCGATGTGACCTGGTTGCAGAAACATGCGCAGCAAGCACTCA 2404
|||||
1690 GluArgSerAsnHisAlaGluAlaIleGlnCysLeuValHisSerAlaAl 1706
|||||
2405 GAACGAACCAATTCATGCTGAAAGCTGCACAGTGTCTACTCCACTCAGCAGC 2454
|||||
1706 alaValAlaGluTyLeuSerMetLeuGluAspArgLysTyLeuProV 1723
|||||
2455 ACTTGTTGCTGAATATTTGAGCATGCTGAGAGACCGGAAATATCTTCTCTG 2504
|||||
1723 aGlyCysValThrPheGlnAsnIleSerSerAsnValLeuGluGluSer 1739
|||||
2505 TGGGATGTGTAACTTTACAGATATTTCAATCATATGTTTAAAGAAATCT 2554
|||||
1740 AlaValSerAspAspValValSerProAspGluGluGlyLysSerg 1756
|||||
2555 GCGGTTCAAGTATGTGTGATATCTCCAGATGAAGAAAGCTTTCGCTCGG 2604
|||||
1756 yLysTyThrPheThrGluSerGlyLeuValGlyLeuLeuGlnAlaAla 1773
|||||
2605 AAATACATTTACTAGTCAGCAGACTTGTGGATTACTGGAACAGACAGCTG 2654
|||||
1773 IasSerPheSerMetAlaGlyMetTyGlnAlaValAsnGluValTyTrpLys 1789
|||||
2655 CTTCCTTCTCTATGTGCTGGCATGTATGACAGTAAATGAAGTTTACAAA 2704
|||||
1790 ValLeuIleProIleHisGluAlaAsnArgAspAlaLysLysLeuSerTh 1806
|||||
2705 GTACTTATTTCTATTCAATGAAGCTAATCGGATGCAAAAGAACTATTCAC 2754
|||||
1806 rIleHisGlyLysLeuGlnGluAlaPheSerLysIleValHisGlnSerT 1823
|||||
2755 AATTCATGTAAACTTCAGAGAGCATTCAGCAAAATGTTTCATCAGAGTA 2804
|||||
1823 hrGlyTrpGluArgMetPheGlyThrTyrrPheArgValGlyPheTyrgly 1839
|||||
2805 CTGGCTGGAGCGGATGTTTGGCACCTATTTTCTGTGTTTATGGA 2854
|||||
1840 ThrLysPheGlyAspLeuAspGluGlnGluPheValTyTrpLysGluProAl 1856
|||||
2855 ACCAAGTTCGGGGATTTGGATGAACAAGAAATTTGTTTACAAGAGCCTGC 2904
|||||
1856 aIleThrLysLeuAlaGluIleSerHisArgLeuGluGlyPheTyrglyG 1873
|||||
2905 AATTAACCAAACTTGACAGAGATATCTCACAGATTGGAGGATTTTACGGAG 2954
|||||
1873 LuArgPheGlyGluAspValValGluValIleLysAspSerAsnProVal 1889
|||||
2955 AAAAATTTGAGAGGAGTGTGTTGAAGTAAATCAAAACTCTAATCTCTGTA 3004
|||||
1890 AspLysCysLysLeuAspProAsnLysAlaTyrrIleGlnIleThrTyrrA 1906
|||||
3005 GACAAAGTCTAAATTTACATCTTAACAAGGCAATATATTCAGATTCATGCT 3054
|||||
1906 IgluProTyrrPheAspThrTyrgluMetLysAspArgIleThrTyrrPhea 1923
|||||
```

```

3055 GGAGCCCTACTTTGACACATATGATGAAGACAGATCCTATTTCG 3104
1923 splysasnlyrAsnleuargrArpHemetyrcystrProphetrleu 1939
|||||
3105 ACAAAATTCANCTCTCGTCGATTCATCTACTGTACACCTTACTTTA 3154
1940 AspglyArGAlahlsGlyluLeuHlsGlyluPheLysArglyThrII 1956
|||||
3155 GATGGCCGTCCTGAGGAACTTCATGACAAATTCAAAAGGACCAAT 3204
1956 eleuThrThSerHisAlaPheProtyrIleuLysThrArgValAsnValIT 1973
|||||
3205 TCTGACTACCTCTCATGCTTCTCTTATATATAAACAAGGTCAAATGTC 3254
1973 hrHlsLysGluGluIleIleLeuThrProIleGluValAlaIleGluAsp 1989
|||||
3255 CTGATAAAGAAGATCATCTTTAACAACAAATTCAGTTGCTATTGAGGAC 3304
1990 MetGluLysLysThrGluGluLeuAlaPheAlaThrHisGlnAspProAl 2006
|||||
3305 ATGAGAGAAAAGACACAGAGATTGGCATTTGCAACACATCAGATCCCG 3354
2006 aAspProLysMetLeuGlnMetValleuGlnLysSerValGlyThrThy 2023
|||||
3355 AGACCCCAAAATGCTTCAGATGCTCTCCAGGATCTGTAGGCACACAG 3404
2023 aLAsnGlnGlyProLeuGluValAlaGlnValPheLeuSerGluIlePro 2039
|||||
3405 TGAATCAGGGGCTTGGAAAGTTGCCAGGTTTCTGCTGTAATAAATCT 3454
2040 SerAspProLysLeuPheArgHlsHlsAsnLysLeuArgLysPheLy 2056
|||||
3455 AGTGACCCCAAGCTTTCACAGCATCATATAAATCACTGCACTGCTTAA 3504
2056 sAspPheThrLysArgCysGluAspAlaLeuArgLysAsnLysSerLeuI 2073
|||||
3505 AGATTTTACTAAAGGTGAGAGATGCTTAAAGAAAAAATAAGAGCTTAA 3554
2073 leGlyProValGlnLysGluTyrGlnArgGluLeuGlnLysLeuSerSer 2089
3555 TTGGCCCGGTTCAAAAGAGATATCAAAAGGAATGGGGAAACATATCTCG 3604
2090 Pro 2090
|||||
3605 CCT 3607

seq_name: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AA507373
seq_documentation_block:
ID AA507373 standard; CDNA; 7215 BP.
XX
XX AA507373;
AC
XX
XX 26-SEP-2001 (first entry)
DT
XX
XX Human CDNA encoding CLASP-5.
DE
XX
XX Human; CLASP-5; cadherin-like asymmetry protein; immune gateway;
KW immunogen; antibody; autoimmune disease; rheumatoid arthritis;
KW multiple sclerosis; leukaemia; insulin dependent diabetes mellitis;
KW acquired immunodeficiency syndrome; AIDS; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 112..6048
FT FT /*tag= a
FT FT /product= "CLASP-5"
FT FT primer_bind 1498..1519
FT FT /*tag= b
FT FT /note= "PCR primer HC59s1"
FT FT primer_bind 3087..3113
FT FT /*tag= c

```

```

FT FT /note= "Antisense oligonucleotide AA507384"
FT FT 3087..3113
FT FT /*tag= d
FT FT /note= "Antisense oligonucleotide AA507384"
FT FT primer_bind 3232..3259
FT FT /*tag= e
FT FT /note= "Antisense oligonucleotide AA507384"
FT FT primer_bind complement (3642..3660)
FT FT /*tag= f
FT FT /note= "PCR primer HC5AS10b"
FT FT primer_bind 4728..4756
FT FT /*tag= g
FT FT /note= "Antisense oligonucleotide AA507384"
XX
XX WO200142296-A2.
XX
XX 14-JUN-2001.
XX
XX 13-DEC-2000; 2000WO-US34163.
XX
XX 13-DEC-1999; 99US-0170453.
XX 14-JAN-2000; 2000US-0176195.
XX 14-FEB-2000; 2000US-0182296.
XX 11-APR-2000; 2000US-0196267.
XX 11-APR-2000; 2000US-0196460.
XX 11-APR-2000; 2000US-0196527.
XX 11-APR-2000; 2000US-0196528.
XX 11-APR-2000; 2000US-0547276.
XX 13-OCT-2000; 2000US-0240503.
XX 13-OCT-2000; 2000US-0240508.
XX 13-OCT-2000; 2000US-0240543.
XX 13-OCT-2000; 2000US-0240539.
XX
XX (ARBO-) ARBOR VITA CORP.
XX
XX Lu P, Garman JD, Candia AF;
XX
XX WPI: 2001-367865/38.
XX P-PSDB; AAU04024.
XX
XX CLASP-5 polynucleotides, proteins and antibodies are used to prevent or
XX treat a CLASP-5 mediated disease, such as an autoimmune disease e.g.
XX Rheumatoid arthritis -
XX
XX Claim 1; Fig 6; 188pp; English.
XX
XX The sequence encodes a cadherin-like asymmetry protein, CLASP-5, which
XX is a transmembrane protein of the immune system involved in the formation
XX of the immune gateway. CLASP-5, polynucleotides encoding it and an
XX anti-CLASP-5 antibody are used to prevent or treat a CLASP-5 mediated
XX disease, such as an autoimmune disease caused or exacerbated by
XX increased activity of TH1 cells. These diseases may include Addison's
XX disease, rheumatoid arthritis, dermatitis, Multiple Sclerosis, Systemic
XX lupus Erythematosus and Autoimmune Thyroiditis, inflammatory conditions
XX (e.g. ischaemia-reperfusion), and responses; Leukaemia, acquired
XX immunodeficiency syndrome, AIDS, syphilis, typhoid, Burkitt's lymphoma,
XX CC Grave's disease and insulin dependent diabetes mellitis. CLASP-5 is used
XX for inhibiting an immune response in a cell (T cell or B cell) by
XX interfering with the expression of a CLASP-5 gene in the cell, the
XX ability of a CLASP-5 protein to bind to another cell or the ability of a
XX CLASP-5 protein to bind to another protein. The antibody binding CLASP-5
XX is used to inhibit an immune response is a subject. The polynucleotide is
XX used to detect CLASP-5 expression in cells and for diagnosis of
XX diseases and disorders associated with aberrant expression of CLASP-5.
XX
XX Sequence 7215 BP; 2005 A; 1734 C; 1721 G; 1755 T; 0 other;
SQ

```

```

alignment_scores:
Quality: 5785.50      Length: 2071
Ratio: 3.487          Gaps: 25
Percent Similarity: 80.106      Percent Identity: 55.722

```

```
alignment_block:
US-09-737-246-2 x AAS07373 ..
```

Align seg 1/1 to: AAS07373 from: 1 to: 7215

44 SerHisHisThrThValProLeuthr.....GluAlaIva 55
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|
34 ACCGGCGGTAACTGTGAATCTTTCACAAAGCCCTCAGTTTATACCCCTG 83

55 IAsPProVALAsPLeuGLuAsPYrLeuILeThHisProLeuAlaVaLa 72
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|
84 GAGGCCAAGTGAGACTTTGAAAGAAGCTTGATGACACACCCTAACAGAGCTGG 133

72 sPSeRcLyProLeuArLysPleuILegLPHePrOProAspRIIEgLu 88
||::|::|::|::|::|::|::|::|::|::|::|::|::|
134 ATGTGCAGCTTCCGACAGAGAGCTCGGGAGCTTCACGTGATGACGACTTGGAC 183

89 ValValTYrSerProArgAsPCysArgTrHLeuValSerAlaValProGI 105
||||::|::|::|::|::|::|::|::|::|::|::|::|::|
184 GTGGCTGTACCCGCCAAGGAATGTAGAGCTTTGGACGCCCTCTTTGCCGGA 233

105 uGlUser...gluMetAsPProHisValArgAsPCysIILeArgSerTYrT 121
||||::|::|::|::|::|::|::|::|::|::|::|::|::|
121 hrcLusprTPalaleValIIleArglys..... 130

284 TCCGTGAGTGGCTTAATCGTGAAACCCGAAAAACAAGAGAGTCCAGAATC 333
:::|::|::|::|::|::|::|::|::|::|::|::|::|
131TyHIsyLSylengLIYrThcLYheasnProksnthrLeuAspy 145
:::|::|::|::|::|::|::|::|::|::|::|::|::|
334 TGTCGCTTTAAAAGCACTGGATCT..... 357

145 scgLusLygLUArgGLnLysGLyeuProLySGlnValPhelGUsErAsp 162
::::|::|::|::|::|::|::|::|::|::|::|::|::|
358 ..CGAAAGATTTCACAAAGACCTTCCGAACAACAGCCTTTGAGTCGGAAA 406

162 LuAlaProAspGLyasnSerTYrGlnAsPaSPGlnAsPaSPLeuLyArg 178
:::|::|::|::|::|::|::|::|::|::|::|::|::|
407 CCTTGGAG.....TGCATGAAACCCGCTGCTCAGGACGCCCCCGCAC 450

179 ArgSerMetSerILEasPsrPhrProArgGLySerTPraPalAcysSerII 195
:::|::|::|::|::|::|::|::|::|::|::|::|::|
451 TTAAAGCTGCTGTGCGACGCTGTCTGGAAAGCCCCGCTACCTGCTGTGA 500

195 ePhAsPLeuLySAsnSerLEuProAsPaLaleuLeuProAsnLeuA 212
||||::|::|::|::|::|::|::|::|::|::|::|::|
501 CTTTGACCTCCGACGCTGCACACCTGACAAACGGCTGAAAAACCTCTGCG 550

212 spArTrHrProAsngInGuJIleAsPaVgGlnAsnaSPAsPaSPGlnArglys 228
:::|::|::|::|::|::|::|::|::|::|::|::|::|
551 AGCAAGAGACTGCCGAGGACTTTGAAAGCAAGAACGAGAGGCCCGGAGG 600

229 SerAsnArgHISlySGlneUphEALaleuHisPseRserProAsPGluGI 245
:::|::|::|::|::|::|::|::|::|::|::|::|::|
601 ACCAAGCAGGAGGCGGAGCTCTTGGCTTACCAATCATGTCAGACAGAGA 650

245 uGIUpTroILieGLuArgLeuSERValProAsPIIEProLySGluHisPhEc 262
:::|::|::|::|::|::|::|::|::|::|::|::|::|
651 GGATGCTGTGGGAATACGTCCACGTACCGAATGTGCCAAGAACACCTGG 700

262 LygInArGLEuLeuValIscYSlaUSerLEuLySPheGIIIEgLIle 278
||::|::|::|::|::|::|::|::|::|::|::|::|
701 GCACACAAATATTGGTCAAGTTGCTGACCTTGAAGTTTCGAGATTGGAATT 750

279 GlUPrIlePheAlaSerLEuAlaLeuTYrAsPyAllySGlySLySLy 295
||||::|::|::|::|::|::|::|::|::|::|::|::|
751 GAAGCCCTGTTGCCAGCATGCTCCCTTACGATGTATAAGAAAGAAAAA 800

295 sLIEserGLuasnPherTYrPhAsPLeuAsnSerGLuImetLYsgLYL 312
:::|::|::|::|::|::|::|::|::|::|::|::|::|
801 GATCTCAGAAAATTTCATCTGACCTTAACCTCGACAGCTTCAAAGAT 850

[illegible]


```

5197 GACAGCATGTTACAGAGAT.....CATAGAGAAATTTGGAACTTA 5240
1832 rPheArGValGlyPheTYrGlyThrLysPheGlyLysPheAspGluGln 1849
5241 CTTCCGAGTGTGTTCTTGGATTCAGAAATTTGGGATTTGGATGACAGG 5290
1849 LpHeValTYrLysGluProAlaIleThrLysLeuAlaGluIleSerHis 1865
5291 AGTTGTCTACAAAGAGCCCTCAATACCAAGCTTCGAGATCTCACAT 5340
1866 ArgLeuGluGlyPheTYrGlyLysLysArgPheGlyLysAspValValGlu 1882
5341 AGACTAGAGGCTTTTATGCTCAATGTTTGTGCGAATTTGTGGAAGT 5390
1882 LLeLysAspSerAsnProValAspLysCysLysLeuAspProAsnLysA 1899
5391 GATTAAAGATCTCACTCTCTGTGCAAAAACCAAGTTGATCTTAACAAG 5440
1899 LaTYrIleGlnIleThrTYrValGluProTYrPheAspThrTYrGluMet 1915
5441 CCTACATACAGATCACTTTGTGGAGCCCTACTTGTAGATGAGATGAGATG 5490
1916 LysAspArgIleThrTYrPheAspLysAsnTYrAsnLeuArgArgPheMe 1932
5491 AAAGACAGGGTCCACTTGTAGAAATTTCAACCTCGGAGGTTGAT 5540
1932 TYrCysThrProPheThrLysAspGlyArgAlaHisGlyLysLeuHisG 1949
5541 GTACACACACCCGTTCCACCTGAGGGGCGCTCGGGAGAGCTGCATG 5590
1949 LpGlnPheLysArgLysThrIleLeuThrThrSerHisAlaPheProTYr 1965
5591 ACCAATACAGAGAACACAGCTGACCACTATGACAGCCCTCCCTTAC 5640
1966 IleLysThrArgValAsnValThrHisLysGluGluIleIleLeuThrPr 1982
5641 ATCAAGACACGAGATCAGCGCTCATCCAGAGAGGAGTTGTTTGACACC 5690
1982 OlleGluValAlaIleGluAspMetGlnLysLysThrGlnGluLeuAlaP 1999
5691 GATTGAGTGGCATTGAAAGACATCAAGACAGACCTGCGATTAGCAG 5740
1999 heAlaThrHisGlnAspProAlaAspProLysMetLeuGlnMetValLeu 2015
5741 TTGCCATTAAACGAGAGCGCCTGATGCAAGATGCTTCAGATGCTGCTG 5790
2016 GlnGlySerValGlyThrThrValAsnGlnGlyProLeuGluValAlaGl 2032
5791 CAAGGCTGTGTGGAGCTACTGTAAATCAGGAGCACCTGGAAGTAGCCCA 5840
2032 nValPheLeuSerGluIleProSerAspProLysLeuPheArgHisHisA 2049
5841 AGTGTTTTGGCTGAATTCCTGCTGATCCAAAACCTATATGCACATCCA 5890
2049 snLysLeuArgLeuCysPheLysAspPheThrLysArgCysLysAspAla 2065
5891 ACAAGTTGAGTTATGCTTTAAGCAATTCATCATGAGATGGTGCAACT 5940
2066 LeuAlaGlyLysAsnLysSerLeuIleGlyProValGlnLysGluTYrGlnAr 2082
5941 GTAGAGAAAAACAGCGTCTCATCAAGGAGACCAAGAGGAATATACACA 5990
2082 gGluLeuGlyLys 2086
5991 GGAAGCTCAAAAAG 6003
seq_name: /SIDS1/gcgdata/hold-1/geneseq/geneseqn-emb1/NA2001B.DAT:ABL17736
seq_documentation_block:
ID ABL17736 standard; DNA: 9389 BP.
XX
AC ABL17736;

```

```

XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 4681.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX Drosophila melanogaster.
PN WO200171042-A2.
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US09231.
PF
XX
XX 23-MAR-2000; 2000US-191637P.
PR
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PERK ) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PMD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PR interactions -
XX
XX Claim 1; SEQ ID NO 4681; 21bp + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB101840-AB16175) and the encoded proteins
CC (AB57737-AB572072).
CC
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX Sequence 9389 BP; 2710 A; 1842 C; 1849 G; 2988 T; 0 other;
SQ
alignment_scores:
Quality: 4690.00 Length: 2340
Ratio: 2.886 Gaps: 58
Percent Similarity: 69.444 Percent Identity: 44.188
alignment_block:
US-09-737-246-2 x ABL17736/rev ..
Align seq 1/1 to reverse of: ABL17736 from: 1 to: 9389
4 ArgArgAlaPheAlaGlnLysIleSer..... 12
:::|||||:|||||:|||||:|||||: 12
8277 CAGGCTGCTGTTGTGCAAAAAGTTTCCAACTAAGTTTTCACCTAAAGAA 8228
12 ..... 12
8227 ATAGTTATTTTAAACAAGAACGCCCTTAATGCTTGTATACATATGTA 8178
13 .....ArgThrValAlaAlaGluValArgLys 21
8177 CATATATCAATTTGTAAATCCCTAGCGAGTGGCGTGGGATGTACGAAAA 8128
22 GlnIleSerGlyGlnTYrSerGlySerProGlnLeuLeuLysAsnLeuAs 38
8127 AATGTTTGTGTTGCCATGTATCAAAAGCAATG.....GACCCATC 8087
38 nLeuValGlyAsnIleSerHisHisThrThrValProLeuThr..... 52

```


[illegible]

```

1728 heGlnAsnIleSerSerAsnValLeuGluIuSerAlaValSerAsp 1744
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2908 TTCACGAAATTTCTCCACACACATAAGTAGTCGGCGGTATCGATGAT 2859
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1745 ValValSerProAspGluGluGlyIleCysSerGlyLysTyrPheThrGl 1761
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2858 GTGCTAAGTCCCGCGGAGATGTCTGCTTAAGAAATCATTCATCTCA 2809
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1761 uSerGlyLeuValGlyLeuLeuGluGlnAlaIleAserPheSerMet 1778
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2808 AACTGGGTTGAAGGCTTCTCTGAGAGAGCTTCACAAATCTTTTCAAG 2759
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1778 IaGlyMetTyrGluAlaValAsnGluValTyrLysValLeuIlePro 1794
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2758 CTGGCATGTATGAGCAATGAGCAAGTGTACAAATCTTATACCCATA 2709
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1795 HisGluAlaAsnArgAspAlaLysLysLeuSerThrIleHisGlyLys 1811
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2708 TGGAGGCTACAGAGATTTTCAAAAGCTAAGCAAGTTCATGCGAAAT 2659
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1811 uGlnGluAlaPheSer..... 1816
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2658 GCAGAGGCAATTTAATCGAATATCCCACTACAGTACACAAATATGTG 2609
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1817 .....LysIleValHisGlnSerThrGlyTyrGluArgMet 1828
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2608 AATTTCACACGAGAAATATATACATATTTAAACAGGCTAAGAGAGAT 2559
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1828 tPheGlyThrTyrPheArgValGlyPheTyrGlyThrLysPheGlyAsp 1845
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2558 TTTGGAGACATCTTCTGTTGGCTTCTTAAGCGGAAATTTGGGGAGCT 2509
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1845 euAspGluGlnGluPheValTyrLysGluProAlaIleThrLysLeuAla 1861
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2508 TGGATCAGCAGGAGATTCATTTAAAGACCAACATTTGACGAAAGTGGCC 2459
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1862 GluIleSerHisArgLeu.Glu..... 1868
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2458 GAAATATTTAGTGGCTCAGGTATATATGCAAAATGGAATAAGAA 2409
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1869 .....GlyPheTyrGlyLysArgPhe 1875
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2408 CTAAATCAATTTGTTTCAACATACGTACGTTTACACTGACGATTC 2359
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1876 GlyGluAspValValGluValIleLysAspSerAspProValAspLys 1892
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2358 GGACCGGACTCTGTGCATATCATTTAAAGATTCACCATCCGTTGATTTAA 2309
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1892 sLysLeuAspProAsnLysAlaTyrIleGlnIleThrTyrValGluPro 1909
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2308 TACCTTGATCCGATTAAGGCTTACATTCAAATTAATTAATGTTGAACCT 2259
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1909 yRPheAspThrTyrGluMetLysAspArgIleThrTyrPheAspLysAsn 1925
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2258 ACTTTGAAACATATGAATGCGTCATCGTAGACATACTTTGAGCGGAA 2209
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1926 TyrAsnLeu..... 1928
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2208 TTCAAATTAAGTATGATATGAATTAACAGATATTTAAATCGAAATTTA 2159
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1929 .....Arg.ArgPheMetTyrCys 1934
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2158 AAATTAATTAATTAACATTTTGGTTATGACAGACGTTTATATATGCG 2109
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1935 ThrProPheThrLeuAspGlyArgAlaHisGlyGluLeuHisGluInP 1951
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2108 ACCGCTTTTACTTAAATAAGGACACATGCGGACACTTAATATGACGAG 2059
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1951 eLysArgLysThrIleLeuThrThrSerHisAlaPheProTyrIleLys 1968
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2058 TAAACGAAACGATTTTACACAGCGGCAATCTTTCCATTAAGTAAATA 2009

```

```

1968 hrArgValAsnValThrHisLysGluGluIleIleLeuThrProIleGlu 1984
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2008 CTCGATTAATGTAATTAAGCCAGACGAAATATGCTGGAACGATGAA 1959
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1985 ValAlaIleGlu..... 1988
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1958 GTGGCAATGTGAAGTAAGTATGTGTAAATCCCAAAAGTTAAATTAAT 1909
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1989 .....AspMetGlnLysTyrThrGlnIleValAlaPhe 2000
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1908 AAAGCTATTCCCTAGAGATATTCAAAAAACAATTTGGAATTTGGCAGCTG 1859
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2000 IaThrHisGlnAspProAlaAspProLysMetLeuGlnMetValLeuGln 2016
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1858 CCAGAAATCAAGAGCCAGACAGACCCCAAAATATTTGCAATATGCTCAG 1809
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2017 GlySerValGlyThrThrValAsnGlnGlyProLeuGluValAlaGlnVa 2033
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1808 GGATGTATGGAAACACCTGTTAACCAAGGACCGATGGAAATGGCGAGCTGT 1759
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2033 lPheLeuSerGluIleProSerAspProLysLeu...PheArgHisHis 2049
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1758 GTTCCTTTCCAAATTTATCCGACGAGACACGTGACGACAAACACCAA 1709
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2049 snLysLeuArgLeuCysPheLysAspPheThrLysArgCysGluAspAla 2065
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1708 ACAAACTCGGTTGTGCTTCCGAGATTTTCAAGCGTTGTGCGTGAAGCT 1659
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2066 LeuArgLysAsnLysSerLeuIleGlyProValGlnLysGluTyrGlnArg 2082
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1658 TTTGAAAAAGAAATCGCAATCATATCTTCAGATCAAAAAGATTAACAA 1609
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2082 GGUleu 2084
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1608 AGAAGCTG 1602
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

seq_name: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAH41952
seq_documentation_block:
ID   AAH41952 standard; DNA; 3173 BP.
XX
AC   AAH41952;
XX
DT   30-AUG-2001 (first entry)
XX
DE   CLASP-3 nucleotide fragment.
XX
KW   Human; cadherin-like asymmetry protein; CLASP; CLASP-3; immune response;
KW   cell surface molecule; transmembrane protein; immunosuppressive; vaccine;
KW   antiinflammatory; antiarthritic; antianemic; dermatological; uropathic;
KW   ophthalmological; antirheumatic; nephrotropic; antithyroid; antidiabetic;
KW   neuroprotective; antistatic; antibacterial; antiseptic therapy;
KW   gene therapy; chromosome 1p31.1; ds.
XX
OS   Unidentified.
XX
PN   WO200142297-A2.
XX
PD   14-JUN-2001.
XX
PE   13-DEC-2000; 2000WO-US34171.
XX
PR   13-DEC-1999; 99US-0170453.
PR   14-JAN-2000; 2000US-0176195.
PR   14-FEB-2000; 2000US-0182296.
PR   11-APR-2000; 2000US-0196267.
PR   11-APR-2000; 2000US-0196460.
PR   11-APR-2000; 2000US-0196527.
PR   11-APR-2000; 2000US-0196528.
PR   11-APR-2000; 2000US-0547276.
PR   13-OCT-2000; 2000US-0240503.
PR   13-OCT-2000; 2000US-0240508.
XX

```

1321	eraIaGeIuSerSerArGSerLeuLeuIIeCySLeuLeuTrpValLeuLys	13337
404	CACACAGAACTCAACTGCAGAACCTTTGGATCTGCTACTTTGGGCTCCAA	453
1338	AsnAlaAspOluThrValLeuGluInuStrPrpMetAspLeuSerValLe	1354
454	AAATGCAGATGAAACAGCTTCTACAGAAAGTGTTTACAGATCTCTCAGTGT	503
1354	uGluLeuAsnArgLeuLeuAspLeuLeuTyrlLeuCySValSerCySPhag	13711
504	GCAGCTAAACCCGGCTATAGATCTGCTTATCTGTGTGCTTCTTGG	553
1371	IuTyrlTySglTySlySValPrheGluArgMeAsnSerLeuThrPhelys	1387
554	AGATATAAGGAAAAAGTTTGAAACGATATAAAGCTTGACCTTTAG	603
1388	LysSerLysAspMetArgAlaLysLeuGluGluAlaIleLeuGlySerI	1404
604	AAATCAAAAACACTACAGAGACCAAGCTTGAGAGACTATTCTTGGAGCAT	653
1404	egIyAlaArgGluGluMetValArgArgSerArgIyGluLeuGluArgS	14211
654	AGTGCCACAGCAAGAAATGCTACGGCGAAGCCGAGACAGCTCGAGAA	703
1421	erProSerGlySerAlaPrheGlySerGluGluAsnLeuArgTrpArgLys	1437
704	GCCCATCTGGAAAGTGCCTTTGGAAAGTCAAGAAATTTAGGTGGAGGAA	753
1438	AspMetThrIstTrpArgIlnAsnTrpGluTySLeuAspLysSerArgAl	1454
754	GATATACCTACCTGGCGGTCAAAACACAGAAAGCTTGACMAATCAAGAC	803
1454	agIuIIeGluHISgluAlaLeuIleAspGlyAsnLeuAlaThrGluAla	14711
804	AGGATTTGAACACGAAAGCTGATTTGATGGAAGCAACTGGCTACAGACAA	853
1471	snLeuIlelleLeuAspTrpLeuGluIleValGlnThrValSerVal	1487
854	AGGATATCATTTTGTAGTATACATTAGAGATTGTGTTCACACCGTTTCTGA	903
1488	ThrGluSerLysGluSerIleLeuGlyGluValLeuLysValLeuLeuH	1504
904	ACGGATCCAAAGAGAGACATCTTGTGTGAGAGTCAAAAGTGTACTAC	953
1504	sSerMetAlaCySasnGlnSerAlaValTyrlLeuGlnHISCySPheat	15211
954	CACCATGCGCTTAAACCAAGAGCTGATTACTACAAACACTGTTTGCTA	10030
1521	hrgIlnArgAlaIleuValSerTySPherProGluLeuLeuPrheGluGluG	1537
1004	CACAGAGACCTTGTTGTTTCAAGATTTCCTGAACCTTTATTGGAAGAAAG	10533
1538	ThrGluGlnCySAlaAspLeuCySLeuArgLeuLeuArgHISCySSe	1554
1054	ACAGACACTGTGTGATTTATAGCTTCAGCGCTTCCACACTGTAGCAG	11033
1554	rSerIleGlyThrIleArgSerHISProSerAlaSerLeuTyrlLeuLeuM	15711
1104	TAGCATCGGTACAAATACGCTACACCCAGAGTGCCTTACTACTATA	11533
1571	erArgGlnAsnPhegIuIleGlyAsnAsnPrheAlaArgValLysMeGln	1587
1154	TGAGGCAAAACTTTAGATTGGGAATTACTTGGCAGGGGTTAAATATGAG	12033
1588	ValProMetSerLeuSerSerLeuValGlyThrSerGlnAsnPhaAsnG	1604
1204	GTACCAATGTCACTATCCCTTGTTGGTGGCACATCTCGAATTTTAAATGA	1253
1604	uGluPrheLeuArgArgSerLeuLysThrIleLeuThrTyrlaGluGlu	1621
1254	AGGATTTCTTAAAGACGTCTCTTAAAGACTATATTTGACATATCTGAAGAG	1303

1621 spLeuGluLeuArgGluThrThrPheProAspGluValGlnAspLeuVal 1637
 1304 ATCTGGAATTTGAGGAAACAAACATTTCTGATCAGGTCCAGATCTGGTT 1353
 1638 PheAsnLeuHisMetIleLeuSerAspThrValLysMetLysGluHisGln 1654
 1354 TTCATCTCTCATATATGATCTTCTGATACCTGAAATGAAGAACACCA 1403
 1654 nGluAspProGluMetLeuIleAspLeuMetTyrArgIleAlaLysGlyT 1671
 1404 GGAGGATCTCTGAAATGTTGATGATCTAATGTACGAAATGGCAAGGTT 1453
 1671 yGlnThrSerProGlu...ArgLeuThrTrpLeuGlnAsnMetAlaGly 1686
 1454 ACCAGACCTCTCCAGATCTGCAGATTGACCTGGTTCAGAACATGGCAGGC 1503
 1687 LysHisSerGluArgSerAsnHisAlaGluAlaGlnCysLeuValHis 1703
 1504 AAGCACTCGAAGCAGAACATCATGTCTGAAGCTGCACAGTCTGTAGTCCA 1553
 1703 sSerAlaAlaLeuValAlaGluTyrLeuSerMetLeuGlnAspArgLysT 1720
 1554 CTCAGCAGACCTGTTGCTGGAATATTGACATGCTGCGAGACCGGAAT 1603
 1720 yTrpLeuProValGlyCysValThrPheGlnAsnIleSerSerAsnValLeu 1736
 1604 ATCTCTCTGCCATGTTGTAACATTTCAAGATATTTCATCTAATGTTTGA 1653
 1737 GluGluSerAlaValSerAspAspValValSerProAspGluGluGlyT 1753
 1654 GAAGATTCGCGCTCAGATGATGTGATCTCCAGATGAAGAAAGGATAT 1703
 1753 eCysSerGlyLysTyrPheThrGluSerGlyLeuValGlyLeuGluG 1770
 1704 CTGCTCTGGAATAATCTTACTGAGTCAGACGACTTGGGATTACTGGAAAC 1753
 1770 IValAlaAlaSerPheSerMetAlaGlyMetTyrGluAlaValAsnGlu 1786
 1754 AAGCAGCTCTCTCTCTATGCTGATGCTGATGAAGCAGTAAATGA 1803
 1787 ValTyrLysValLeuIleProIleHisGluAlaAsnArgAspAlaLys 1803
 1804 GTTTACAAAGTACTTATTCCTATTCATGAGCAAGCTAATCGGAGTGAAGAA 1853
 1803 sLeuSerThrIleHisGlyLysLeuGlnGluAlaPheSerLysIleValH 1820
 1854 ACTATCCAAATTCATGTGTAACCTTCAGAACATTCACCAAAATTTGTTTC 1903
 1820 IsGlnSerThrGlyTyrGluArgMetPheGlyThrTyrPheArgValGly 1836
 1904 ATCAGAGTACTGCTGGTGGAGCGAGATGTTGGCAGCTATTTCTGTGGT 1953
 1837 PheTyrGlyThrLysPheGlyAspLeuAspGluGlnGluPheValTyrLys 1853
 1954 TTTTATGGAACCAAGTTCGGGATTTGATGAACAAGAAATTTGTTTACAA 2003
 1853 sGluProAlaIleThrLysLeuAlaGluIleSerHisArgLeuGluGlyP 1870
 2004 GGAAGCTCGCAATACCAACTTCGACAGATATCTCAGATTTGGAGGAT 2053
 1870 heTyrGlyGluArgPheGlyGluAspValValGluValIleLysAspSer 1886
 2054 TTTACGAGAGAAAGATTGGAGAGATGTGTTGAAGTATTCAAAGCTCT 2103
 1887 AsnProValAspLysCysLysLeuAspProAsnLysAlaTyrIleGlnI 1903
 2104 AATCTGTGAGACAAAGTAAATTAGATCTTAACAGGCAATATATTCAGAT 2153
 1903 eThrTyrValGluProTyrPheAspThrTyrGluMetLysAspArgIleT 1920
 2154 TACCTATGTGGAGCCATCTTGTGACACATATGAGATGAAGGACAGAAATCA 2203
 1920 hrTyrPheAspLysAsnTyrAsnLeuArgArgPheMetTyrCysThrPro 1936

2204 CCTATTTCGACAAATAATACAACTTCGTGATTCATGTACTGTACACCC 2253
 1937 PheThrLeuAspGlyArgAlaHisGlyGluLeuHisGlnGlnPheLysAr 1953
 2254 TTTACTTTAGATGGCGGTGCCATGGGGAAGCTTCATGCAACAAATTCAAAAG 2303
 1953 GlyThrIleLeuThrThrSerHisAlaPheProTyrIleLysThrArgy 1970
 2304 GAACACCATTCGTGACTACGTCTCATGCTTCTCTATATTAACAAAGG 2353
 1970 AlaAsnValThrHisLysGluGluIleIleLeuThrProIleGluValAla 1986
 2354 TCAATGTACATCAATAAGAGAGATCATTTTAACACCAATTTGAAGTTGCT 2403
 1987 IleGluAspMetGlnLysLysThrGlnGluLeuAlaPheAlaThrHisGln 2003
 2404 ATTGAGACATGCGAAGAAAGACACAGAGAGTTGGCAATTTGCAACACATCA 2453
 2003 nAspProAlaAspProLysMetLeuGlnMetValIleuGlnGlySerValG 2020
 2454 GGATCCCGCAGACCCCAAAATGCTTCAGATGTACTCCAGGATCTGTAG 2503
 2020 LyrThrThrValAsnGlnGlyProLeuGluValAlaGlnValPheLeuSer 2036
 2504 GCACACAGATGAATCAGAGGCGCTTGGAGAGTTGCCAGTTTTCGTCTCT 2553
 2037 GluIleProSerAspProLysLeuPheArgHisHisAsnLysLeuArgyle 2053
 2554 GAAATACCTAGTACCCAAAGCTCTTCAGACATCATATTAATGCTGCACT 2603
 2053 uCysPheLysAspPheThrLysArgCysGluAspAlaLeuArgLysAsnL 2070
 2604 CTGCTTTAAATATTTTACTAAAGGTGTGAAGATGCTTAAGAAAAATA 2653
 2070 ySerLeuIleGlyProValGlnLysGluTyrGluArgGluLeuGlyLys 2086
 2654 AGACCTTAATTTGGCCGCTTCAAAAGAGATATCAAAAGGAATTTGGGAAA 2703
 2087 LeuSerSerPro 2090
 2704 CTATCTTCGCT 2715

seq_name: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT:ABLI7737
 seq_documentation_block:
 ID ABLI7737 standard; DNA; 5589 BP.
 AC ABLI7737;
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 4684.
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.
 XX
 OS Drosophila melanogaster.
 XX
 PN W0200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NV.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.

144 pLysGlnLysGluArgGlnLysGlyLeuProLysGlnValPheGluSerA 161

541	GCAGGAATTCGAGATCGACAAAGAT.....	TTTTGGTCAT	575
161	spgUalAProAspGlyAsnSerTyrGlnAspAspGln..	AspAspLeu	176
576	TTGAA.....GAGCGCTGGCTTATAAAGGAGAGCGCTGCATCCAGT		619
177	LysArgArgSerMet.....	SerLeuAsp..	185
620	TCTGCACATCATTCGAAATTTGGCTTCAGTAACTCTTGTACCGACAC		669
186	...ThirProArgGlySerThrAlaCysSerLeuPheAspLeuLysAsn		201
670	TTTGACCCCCGGGTGTTTGGCGTAC.....TTGATCTTGCACGTT		713
201	eTLeuProAspAlaLeuLeuProAsnLeuLeuAspArgThirProAsnGln		217
714	CGGTTAAAGACCCCAATTAATACCCATTTGGCTTGACAAATGTAACCTCCGAG		763
218	GluIleAspArgGlnAsnAspAspGlnArgLysSerAsnArgHisLysGln		234
764	CATATGTGATCAATCCAAATGAAGACCGCGCAGCAAGATGCGCAATGAGC		813
234	uLeuPheAlaLeuHisProSerProAspGlnGlnGluProIleGluArgL		251
814	TCCTTTTCACCTGTAATCCGGAATCTGAAGCGGAGGAAAAACATTTGAACGC		863
251	euserValProAspArgProLysGlnHisPheGlnGlnArgLeuLeuVal		267
864	GCCTCTCGCAGAGACTTCAGTGTGACATATGGTCATCGCATTCAGGTT		913
268	LysCysLeuSerLeuLysPheGlnIleGluIleGluProIlePheAlaSe		284
914	AACTCCCTACAGCTTCGGCTTGACGCTGGAGGTGAGCGCAATTTTGGCATC		963
284	rLeuAlaLeuLyrAspValLysGlnLysLysLysIleSerGlnAsnPhet		301
964	GATGGCTATTATATGCTAAAGGAACGGCAAAAAATTTCTGAAAAATTTCT		1013
301	YrPheAspLeuAsnSerGlnGlnMetLysGlnLeuLeuArgProHisVal		317
1014	ATTTTGACATGACCCAGCATAGTACTTTAAACGCATGTTTGCATGATCATGTA		1066
318	ProProAlaAlaIleThrThrLeuAlaArgSerAlaIlePheSerIleTh		334
1064	CAGTGGCCGCAATGATGAATACCGACAGCCACATTCAGCATATTCGAAATTTTC		1114
334	rTyrProSerGlnAspValPheLeuValIleLysLeuGlnLysValLeuG		351
1114	ATATCCGAGTAACGATTTATTTCTCGTATTCGTCGTGAAAAAGCTTTTG.		1162
351	IngIleValAspIleGlyLysAlaGlnArgProTyrMetIlePheLysGln		367
1163	..CAAGGTGACATTAACAATTCCTCGTCGAGCCGTATCTC.....		1198
368	AlaAspAlaThrLysAsnLysGlnLysLeu...GlnLysLeuLysSerGln		383
1199AAAGAAAGACAAGATTAATATACGAGAGAAAGTAATAATCCAA		1239
383	nAlaAspGlnPheCysGlnArgLeuGlyLysTyrArgMetProPheAlaT		400
1240	CGCTGTGACATCACTGTGTAACACATAGGCAAAATACCGAATGCCGTTTGAT		1288
400	rPThrAlaIleHisLeuMetAsnIleValSer.....		410
1290	GGACAGCAATCTACCTTAACAATGTTTTTAATGGCGATTAACCTTGAAGAC		1339
411	...SerAlaGlySerLeuGlnArgAspSer.....		419
1340	AAAGACGCTGGCGCGCGGAGACGGGATTCCTTTGGAAATGCTTGTCGAG		1389
420ThrGluValGluIleSerThrGlyGlnArgLysLysSerT		433

977 TCATCCGTGATATGCGAAGAAACATTATCCAGGCGACCTCGGCTCGAC 1026
1490 SerLysGluSerLLeuGluGlyValLeuLysValLeuLeuHisSerIle
1027 TGTAAAGACAGCCTGCTGGAGAGTGTCTGAGGAGCTGCTGGTAATTCCT
1506 TAlaCysAngInSerAlaValTyrLeuGlnHisCysPheAlaThrGlnA
1077 GAACCTGTGATCAGAGTACACCTACCTGACTGACTGCTTTCACACACTCC
1523 rAlaLeuValSerLysPheProGluLeuLeuPheGluGluGluThrGlu
1127 GTCCTCTCATCGCCAAAGTTTGGAGACTTACTCTTCGAAAGAGAGGTGAA
1540 GlnCysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSerL
1177 CAGTGTTCAGCTATGTACACAGTCTGACACACTGACAGCAGCAGCAT
1556 eGlyThrIleArgSerHisProSerAlaSerLeuTyrLeuLeuMetArg
1227 GGATGTACCCGAGCGCAAGCCTGTGCCACCTTAACTCTCATAGAGGT
1573 LAsnPheGluIleGlyAsnAsnPheAlaArgValLysMetGlnValPro
1277 TCAGTTTGGAGCCACCAAGTAATTTTGCAGAGTAAGATGCAAGTAACC
1590 MetSerLeuSerSerLeuValGlyThrSerGlnAsnPheAsnGluGluP
1327 ATGTCCCTGGCATCTTGTGGGAAAGACACACACTTAAATGAAGAGA
1606 eLeuArgArgSerLeuLysThrIleLeuThrTyrAlaGluAspLeuG
1377 CCGAAGAAATCTTGAGGACAATTTTGGCTTATCAGAAAGAGACACAG
1623 LuleuArgGluThrThrPheProAspGlnValGlnAspLeuValPheAsn
1427 CCATGCGAGTGAATCCTTTCCACCCAGGTGAGGAACTCTCTGTAT
1640 LeuHisMetIleLeuSerAspThrValLysMetLysGlnHisGlnGluAs
1477 CTGAATAGCATCTTATATGACACAGTGAATAATGAGGAAATTTCAGGAAGA
1656 ProGluMetLeuIleAspLeuMetTyrArgIleAlaLysGlyTyrGluT
1527 TCCTGAGATGCTTATGATCATGTACAGAAATGGCAAGATGTACCAAG
1673 hSerProGlu...ArgLeuThrThrPheGlnAsnMetAlaGlyLysHis
1577 CATCTCTGATCTCGCGCTGACCTGGCTCCAGAACTGGCAGAGAAACAC
1689 SerGluArgSerAsnHisAlaGluAlaAlaGlnCysLeuValHisSerAl
1627 ACCAAGAAAGAGTCTACACGAGGCTGCCATGTGCTGTGCACGCCCC
1705 aAlaLeuValAlaGluTyrLeuSerMetLeuGluAspArgLysTyrLeuP
1677 TCGGTAGTGGCTGAGTATCTGAGCATGCTGAGAGCACACACTACCTGC
1722 roValGlyCysValThrPheGlnAsnIleSerSerAsnValLeuGluGlu
1727 CCGTGGGCGAGTGCCTTCAGAAATATTTCTTCATGTGCTGGAGGAG
1739 SerAlaValSerAspAspValValSerProAspGluGluGlyIleCysSe
1777 TCTGTGTCTCTGAGACACCCCTGTACCTGACAGAGATGGGCTGTGCCCC
1755 rGlyLysTyrPheThrGluSerGlyLeuValGlyLeuLeuGlnAlaAla
1827 AGCCGAGTACTTCACCGAGAGTGGCTGTAGGCCCTCGAGACAGGCCG
1772 LAlaSerPheSerMetAlaGlyMetTyrGluAlaValAsnGluValTyr
1877 CCGAGCTTTCAGCACGGAGGCTTATATGAGCAGTTAATGAGGTCTAC 1926

1789 LysValLeuIleProIleHisGluAlaAsnArgAspAlaLysLysLeuSe
1927 AAGCTGGCATCCCATCTCTAGAACGCGCATCGAAGTAATTCGGAAGCTGAC
1805 rThrIleHisGlyLysLeuGlnGluAlaPheSerLysIleValHisGlnS
1977 ACTGACTCAGCAAGCTGCAGAGAGCCCTTCGACAGCATGTTAAACAAG
1822 eThrGlyTyrGluArgMetPheGlyThrTyrPheArgValGlyPheTyr
2027 AT.....CATAAAGAAATGTTTGAACCTTACTTCCAGTGGTGTCTT
1839 GlyThrLysPheGlyAspLeuAspGluGlnGluPheValTyrLysGluP
2071 GGATCCAAATTTGGGGAATTTGATGAACAGAGAGTTTCTACAAAGAGCC
1855 aAlaIleThrLysLeuAlaGluIleSerHisArgLeuGluGlyPheTyrG
2121 TGCATTAACCAAGCTCTCTGATCTCATAGACTAGACATTTTATG
1872 LysGluArgPheGlyGluAspValValGluValIleLysAspSerAspPro
2171 GTCAATGTTTGGTGCAGAAATTTGTGAAGTGAATTAAGACTCCACTCT
1889 ValAspLysCysLysLeuAspProAsnLysAlaTyrIleGlnIleThrTy
2221 GTGACAAAGAACCAAGTTGGATCTTACACAGGCCCTACATACGATCATTT
1905 rValGluProTyrPheAspThrTyrGluMetLysAspArgIleThrTyr
2271 TGTGAGCGCCCTTCTGTGATGATGAGTAAGAAAGACAGGCTCAATACCT
1922 hAspLysAsnTyrAsnLeuArgArgPheMetLysCysThrProPheThr
2321 TTGAGAAATTTCAACCTCCGAGGTTTCATGTACACACCCGCTGAC
1939 LeuAspGlyArgAlaHisGlyGluLeuHisGluGlnPheLysArgLysThr
2371 CTGAGGAGGCGGCTCGGGGAGAGCTGCATGACAGTACAGAAAGAACAC
1955 rIleLeuThrThrSerHisAlaPheProTyrIleLysThrArgValAsnY
2421 AGTCTGACCACTATGACAGCGCTCCCTACATCAACACAGAGTACGCG
1972 aThrHisLysGluGluIleLeuThrProIleGluValAlaIleGlu
2471 TCATCCAGAGAGAGAGTGTGTTTGAACCCGATTGAGATGGCCATTGAA
1989 AspMetGlnLysLysThrGlnGluLeuAlaPheAlaThrHisGlnAspPr
2521 GACATGAAGAAAGAACCCCTGCACTTAGCAGTTCCTTAAACAGAGGCC
2005 aAlaAspProLysMetLeuGlnMetValLeuGlnGlySerValGlyThrT
2571 GCCCTGATCAAAAGATGCTTCAGATGCTGCTGCAAGGCTGTGTGGAGCTA
2022 hValAsnGlnGlyProLeuGluValAlaGlnValPheLeuSerGluIle
2621 CTGTAAATCAGAGGACACTGGAACTACGCCCAAGTGTGCTGGCAATTT
2039 ProSerAspProLysLeuPheArgHisHisAsnLysLeuArgLeuCysPh
2671 CCTGCTGATCCAAACCTCTATCGACATCACACAAAGTTGAGGTATATCTT
2055 eLysAspPheThrLysArgCysGluAspAlaLeuArgLysAsnLysSerL
2721 TAAAGAAATTCATCATGATGTGTGAAGCTGTGAGAAAGAAACAAAGCGTC
2072 euIleGlyProValGlnGluGlyTyrGlnArgGluLeuGluGlyLys
2771 TCATTCACGCGACACCGAGGAGGAAATATCAGCAGGAATCTCAAAAAG 2814

```

seq_name: /SISL/gcdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAH43850
seq_documentation_block:
ID   AAH43850 standard; cDNA; 2148 bp.
XX
XX   AAH43850;
XX
DT   04-SEP-2001 (first entry)
XX
XX   Preliminary human CLASP-7 encoding cDNA sequence.
XX
KW   Human: CLASP-7; cadherin-like asymmetry protein; immune response;
KW   neuroprotective; antidiabetic; immunosuppressive; antirheumatic;
KW   antiarthritic; hypotensive; anti-HIV; cytostatic; immunostimulant;
KW   antiaiematic; antiinflammatory; ophthalmological; nephrotrophic;
KW   antihypertoid; antiasthmatic; antiallergic; antibacterial; gene therapy;
KW   chromosome 19q13.2; autoimmune disease; multiple sclerosis; toxemia;
KW   juvenile diabetes; rheumatoid arthritis; pruritic urticarial papule;
KW   hypertension; Rh incompatibility; ss.
XX
XX   Homo sapiens.
XX
FH   Key Location/Qualifiers
FT   CDS 1..1932
FT     /tag="a
FT     /partial
FT     /product="Preliminary CLASP-7 protein"
FT     /note="no start codon given"
XX
XX   MO200142295-A2.
XX
XX   14-JUN-2001.
XX
XX   PF 13-DEC-2000; 2000MO-US34152.
XX
XX   PR 13-DEC-1999; 99US-0170453.
XX   PR 14-JAN-2000; 2000US-0176195.
XX   PR 14-FEB-2000; 2000US-0182296.
XX   PR 11-APR-2000; 2000US-0196267.
XX   PR 11-APR-2000; 2000US-0196460.
XX   PR 11-APR-2000; 2000US-0196527.
XX   PR 11-APR-2000; 2000US-0196528.
XX   PR 11-APR-2000; 2000US-0547276.
XX   PR 13-OCT-2000; 2000US-0240503.
XX   PR 13-OCT-2000; 2000US-0240508.
XX   PR 13-OCT-2000; 2000US-0240539.
XX   PR 13-OCT-2000; 2000US-0240543.
XX
XX   (ARBO-) ARBOR VITA CORP.
XX
XX   Lu P, Garman JD, Candia AF;
XX
XX   WPI: 2001-381641/40.
XX
XX   DR P-PSDB: AAB99540.
XX
XX
XX   Novel cadherin-like asymmetry protein-7 and polynucleotides encoding
XX   the polypeptide, useful for treating autoimmune disease,
XX   hypersensitivity, preventing transplant rejection by modulating immune
XX   response
XX
XX   Example 3; Fig 1; 151pp; English.
XX
XX   The present invention describes a human cadherin-like asymmetry protein
XX   (CLASP), designated CLASP-7. The CLASP-7 protein (I) and its encoding
XX   nucleotide sequence (II) have activities including: neuroprotective;
XX   antidiabetic; immunosuppressive; antirheumatic; antiarthritic; anti-HIV;
XX   hypotensive; cytostatic; immunostimulant; antiaiematic; antiinflammatory;
XX   ophthalmological; nephrotrophic; antihypertoid; antiasthmatic;
XX   antiallergic; and antibacterial. (II) and CLASP-7 antibodies (III) are
XX   useful for detecting the CLASP-7 polypeptide. (II) is useful for
XX   producing (I) by recombinant methods. (I) or its fragment are useful for
XX   inhibiting an immune response in a cell such as T cell or B cell. A
XX   pharmaceutical composition (C), comprising (I) or (II), can be useful
XX

```

for treating CLASP-7-mediated disease such as an autoimmune disease
CC caused or exacerbated by increased activity of T helper cells. Autoimmune
CC diseases which can be treated using (C) include multiple sclerosis,
CC juvenile diabetes and rheumatoid arthritis. (I) is useful for treating
CC toxæmia or pregnancy induced hypertension, pruritic urticarial papules
CC and Rh incompatibility. (I) is also useful as a diagnostic reagent for
CC immune and other disorders, since diseases characterised by
CC overproduction or depletion of lymphocytes in blood or other organs may
CC be detected by monitoring the level of (I) or its mRNA. CLASP-7 has been
CC mapped to the chromosomal location 19q13.2. The present sequence
CC represents the preliminary human CLASP-7 encoding cDNA sequence which is
CC given in the present invention.

XX

SQ Sequence 2148 BP; 468 A; 649 C; 625 G; 406 T; 0 other:

alignment_scores:
Quality: 2552.50 Length: 599
Ratio: 4.538 Gaps: 1
Percent Similarity: 93.990 Percent Identity: 80.467

alignment_block:
US-09-737-246-2 x AAHA3850 ..

Align seg 1/1 to: AAHA3850 from: 1 to: 2148

1487 ValThrGluSerLysGluSerIleLeuGIlyValLeuLysValLeuLe 1503
:::|||||:::|||||:::|||||:::|||||:::|||||
1 CTTTCAGAAAGCCCCGGAGACGCCTTGTGGGGGATTTCGTGAAGTTTGCT 50
UHisSerMetAlaCysAsnGlnSerAlaValTYrLeuGlnHisCysPheA 1520
|-:-|||:::|||||:::|||||:::|||||:::|||||
51 GTACAGCCTGGGCGAGTGGCCAGAGTGGCCCTTCTTCACATCAGGCCCTGG 100
1520 IaThrGlArgAlaLeuValSerLysSphProGluLeuLeuPheGluGlu 1536
|||||:::|||||:::|||||:::|||||:::|||||
101 CCACCAGAGGGCCCTTGTGTCAAGTCCCGGAGCGCTGTTCGGAGAG 150
1537 GluThrGluGlnCysAlaAspLeuCysLeuArgLeuLeuArgHisCysSe 1553
:::|||||:::|||||:::|||||:::|||||:::|||||
151 GACACGAGAGCTGTGTGCCGACTGTGCTTAGAGCTCTCTACGACACGTGTG 200
1553 rSerSerIleGlyThrIleArgSerHisProSerAlaSerLeuTYrLeu 1570
|-:-|||:::|||||:::|||||:::|||||:::|||||
201 CAGCCGATCATCAGACACATCCGACGACAGCAGCGCTCCGTCTACCTCGC 250
1570 euMetArgGlnAsnPhesIuIleGlyLysAsnSnpHeaIlaArgValLysMet 1586
|||||:::|||||:::|||||:::|||||:::|||||
251 TCATCTCGACAGAAGACTTCAGATTCGGCCACAATTGGCCCTGTGAAGATG 300
1587 GlnValPrometSerLeuSerSerLeuValGIYThrSerGlnAsnPheAs 1603
|||||:::|||||:::|||||:::|||||:::|||||
301 CAGGTGCACCATGTCTCTCTGCTCCCTGTGGGAGCACAGCAGAACCTTCAG 350
1603 nGlnIuPheLeuArgArGSerLeuLysTrHileLeuThTYrAlaGluG 1620
|||||:::|||||:::|||||:::|||||:::|||||
351 TGAAAGACACCTCGAGCATCTCATAAACAACATCTCACCATTAGTGAGG 400
1620 IuAspLeuGluLeuArgGluThrThrPheProAspGlnValGlnAspLeu 1636
|||||:::|||||:::|||||:::|||||:::|||||
401 AGGACATGGGGCTGGGGGACGACCTTCCACAGAGGCTCCAGGACCTTG 450
1637 ValPheAsnLeuHisMetIleLeuSerAspThrValLysMetLysGlnH 1653
|||||:::|||||:::|||||:::|||||:::|||||
451 ATGTTCAACCTGCACATGATCTCTACGCGACACGGTGAAGTGAAGAGACA 500
1653 sGlnGluAspProGluMetLeuIleAspLeuMetTYrArgIleAlaLysG 1670
|||||:::|||||:::|||||:::|||||:::|||||
501 CCAGAGAGACCCCTGAGATGCTCATCGACCTTCATGACAGATTGGCCGGG 550
1670 LYTyrGlnThrSerProGlu...ArgLeuThrTrpLeuGlnAsnMetAla 1685
|||||:::|||||:::|||||:::|||||:::|||||
551 GCTACAGAGGCTCAACCGGACCTTGGCTGTGACCTGTGTGCAAGAACATGGCC 600

1686	gLyysHisSerGluuArgSerAsnHisAlaGluAlaAlaGlnCysLeuA	1702
601	GGGAACACACGGGAGACTGGGCAACACGGCGCGCCAGTGCATGCT	650
1702	HisSerAlaAlaLeuValAlaGluTyrLeuSerMetLeuGluAspArgL	1719
651	GCACGGCGCGCCCTCGTGGCGTAATACCTGCCTGCCTGCAGAGCAC	700
1719	ysrTyrLeuProValAlGlyCysValTyrPheGlnAsnLLeSerSerAsnVal	1735
701	GCACCTGCCTCGCGGAGCTGGCGTTCTCTTCACGAACATCTCATCCAAAGTG	750
1736	LeuGluGluSerAlaValSerAspAspValAlSerProAspGluGluG	1752
751	CtAAGAGAGTCCGCCATCTCCAGACACTCTGTCTGCCACAGAGAGCG	800
1752	yLlCysSerGlyLysTyrPheThrGluSerGlyLeuValGlyLeuG	1769
801	CTTCTGCTCCGGGAAGCACTTCACTAGACCTGGCGCTGTAGAGTTGCTCG	850
1769	lughAlaAlaAlaSerPheSerMetAlaGlyMetTyrGluAlaValAsn	1785
851	AACAGGACCGCGGCTACTTCAACCAATGGCGGGGCTCTACGAGCGCGTAAT	900
1786	GluValTyrLysValLeuLLeProLLeHisGluAlaAsnArgAspAlaL	1802
901	GAGGTCTCAAAACCTCATCCATCTCTGGAAAGCCACCGTGTCTCAAA	950
1802	stLysLeuSerThrLLeHisGlyLysLeuGluGluAlaPheSerLlLe	1819
951	GAGCGTGGCGCGGTGCACGCGCAAACTCAGAGAGCGCTTCCACCAAGATCA	1000
1819	alHisGluSerThrLlLlPheGluMetPheGlyThrTyrPheArgVal	1835
1001	TGCACCAAGTTCGCGCTGGAGCGCGTTCGGAGAGTATTCCTCCGGTG	1050
1836	GlyPheTyrGlyThrLysPheGlyAspLeuAspGluGluGluPheValT	1852
1051	GGCTTCTACGGCGCCCACTTCGTGACTCGATGATGACAGAGTTGTGTCA	1100
1852	rLysGluProAlaLlLlThrLysLeuAlaGluLLeSerHisArgLeuG	1869
1101	CAAGGAGCATGATCATCAAGAACTGGCAGAAATCTCAACCGCGTGGAG	1150
1869	lyPheTyrGlyLysLysArgPheGlyLysAspValAlaGluValLlLeLysAsp	1885
1151	AGTTCTACACGGAGAGATTGGCGCGCGCTGTGGAGATATTCAAAGAC	1200
1886	SerAsnProValAspLysCysLysLeuAspProAsnLysAlaTyrLlLeG	1902
1201	TCTTACCTCGTGGCAAGTCCAGGTTTACTCACAAAAGCCTTCAATCCA	1250
1902	nlLlLlTyrValGluProTyrPheAspThrTyrGluMetLysAspArgL	1919
1251	GATCAGATGTGTGGAAACGTTACTTTGATTACTACAGAGTCAAGACCGGG	1300
1919	lLlLlTyrPheAspLysAsnTyrAsnLeuArgPheMetTyrCysThr	1935
1301	TGACCTACTTTTACCGCAACATATGGCGTTTCCACATCTCTTCTTGACGC	1350
1936	ProPheThrLeuAspGlyArgAlaHisGlyGluLeuHisGluGluPheL	1952
1351	CCGTTACCGCGCGGATGGCGCGCACACGGGAGGCTGCCCGAGCAACAA	1400
1952	sArgLysThrLlLeuThrThrSerHisAlaPheProTyrLlLeLysThrA	1969
1401	GGCTAAGAGCGTGTCTCAGACACGACCAAGCGCTTCCTCCATCATCAAGCTC	1450
1969	rgValAsnValThrHisLysGluGluLlLeuThrProLlLeGluVal	1985
1451	GCATCGGTGTGGCCACGGGAGAGAGAGTGTGTGTACGCCCATGTGAAGTG	1500

```

1986  ALaIleGluAspMetGlnLysLysThrcIngluLeuAlaPheAlaThrHi 2002
1501  GCCATGGAGGACATCGACAGAAGAACACAGAGCGCTGGGCTTTGGCACCGA 1550
2002  sgiNaSPProAlaAspProLysMeLeuGluMeValLeuGlnLysSerV 2019
1551  GCGAGGCCAACCAACGATGCTAAGATCTTACAGTGGTGGCTTAGGGGCTCG 1600
2019  aGIuIhThrValAsnGlnGlnGlyProLeuGluValAlaGlnValPheLeu 2035
1601  TAGGGCCCAACCGTGAACCAACGAGGTCCCGCTGGAGGTGGCCGAGTCTTTTA 1650
2036  SerGluIleProSerAspProLysLeuPheArgHisAsnLysLeuAr 2052
1651  GCAGAGATCCCGGAAACCCCAAGCTCTTCGGCATCACACAATTTGGC 1700
2052  gIeouGysPheLysAspPheThrLysArgCysGluAspAlaLeuArGlySa 2069
1701  GCGCTCTCTTCAAGGACTTCGTCMAAATAATGTAGAGATGCCGCTGGGAAAA 1750
2069  snIysSerLeuIleGlyProValGlnLysGluYrYrGlnArgGluLeu 2084
1751  ATTAGGCCCTGATTTGGGCCGGGACCCAGAAAGGAGTACCAACCTGGAGCTG 1797

seq_name: /SID51/gcdata/hold_geneseq/geneseqn_emb1/NA2001A.DAT:AS08334

```

seq_documentation_block:	
ID	AA508334 standard; CDNA: 6454 BP.
XX	
AC	AA508334;
XX	
DT	23-OCT-2001 (first entry)
XX	
DE	Human CDNA encoding CLASP-4.
XX	
KW	Human; CLASP-4; cadherin-like asymmetry protein-4; autoimmune disease;
KW	ataxia telangiectasia; Human immunodeficiency virus infection;
KW	inflammatory disease; rheumatoid arthritis; multiple sclerosis;
KW	diabetes mellitus; immune disorder; Guillain-Barre syndrome;
KW	severe combined immunodeficiency; allergic reaction; asthma;
KW	immunogen; antibody; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
FT	95..6121
FT	/*tag= a
FT	/product= "CLASP-4"
FT	replace (100,T)
FT	/*tag= b
FT	/note= "Missense mutation"
FT	replace (1095,G)
FT	/*tag= c
FT	/note= "Missense mutation"
FT	replace (1351,G)
FT	/*tag= d
FT	/note= "Missense mutation"
FT	replace (1628,G)
FT	/*tag= e
FT	/note= "Changes Thr to Ala"
FT	replace (2013,G)
FT	/*tag= f
FT	/note= "Changes Asn to Ser"
FT	replace (2421,A)
FT	/*tag= g
FT	/note= "Changes Gly to Asp"
FT	replace (2431,G)
FT	/*tag= h
FT	/note= "Missense mutation"
FT	replace (2439,C)
FT	/*tag= i
FT	/note= "Changes Gly to Glu"
FT	replace (2480,T)

FT FT /*tag= j
FT /note= "Changes Ile to Phe"
FT replace (2498,G)
FT /*tag= k
FT /note= "Changes Phe to Val"
FT replace (2649..2655,AA)
FT /*tag= l
FT /note= "Deletion polymorphism, changes
FT Lys-Tyr-Ser-Phe-Arg to Lys-Leu-Pro-Thr-STOP"
FT replace (2812,G)
FT /*tag= m
FT /note= "Missense mutation"
FT replace (2908,C)
FT /*tag= n
FT /note= "Missense mutation"
XX WO200142294-A2.
XX
XX
XX PD 14-JUN-2001.
XX
XX PF 13-DEC-2000; 2000WO-US34151.
XX
XX PR 13-DEC-1999; 990US-0170453.
XX PR 14-JAN-2000; 2000US-0176195.
XX PR 14-FEB-2000; 2000US-0182296.
XX PR 11-APR-2000; 2000US-0196267.
XX PR 11-APR-2000; 2000US-0196460.
XX PR 11-APR-2000; 2000US-0196527.
XX PR 11-APR-2000; 2000US-0196528.
XX PR 11-APR-2000; 2000US-0547276.
XX PR 13-OCT-2000; 2000US-0240503.
XX PR 13-OCT-2000; 2000US-0240508.
XX PR 13-OCT-2000; 2000US-0240543.
XX PR 13-OCT-2000; 2000US-0240539.
XX
XX (ARBO-) ARBOR VITA CORP.
XX
XX
XX PI Lu P, Garman JD, Candia AF;
XX
XX WPI: 2001-367864/38.
XX P-PSDB; AAU04090.
XX
XX
XX PT Isolated cadherin-like asymmetry protein (CLASP) 4 useful for treating
XX and preventing immune disorders such as arthritis, diabetes and allergy
XX
XX
XX PS Claim 1; Fig 5A; 172pp; English.
XX
XX The sequence encodes human CLASP-4 (cadherin-like asymmetry protein-4)
XX which is a transmembrane protein of the immune system involved in the
XX formation of the immune gateway. CLASP-4, polynucleotides encoding it and
XX an anti-CLASP-4 antibody are used to prevent or treat a CLASP-4 mediated
XX disease, such as an autoimmune disease caused or exacerbated by increased
XX activity of Th1 cells. These diseases may include blood protein
XX disorders, ataxia telangiectasia, common variable immunodeficiency,
XX diGeorge syndrome, HIV infection, leukocyte adhesion deficiency syndrome,
XX lymphopoenia, phagocyte bactericidal dysfunction, severe combined
XX immunodeficiency, anaemia, Wiskott-Aldrich disorder, thrombocytopoenia,
XX haemoglobinuria, allergic reactions such as asthma, organ rejection,
XX graft versus host disease, sepsis, and inflammation such as arthritis,
XX inflammatory bowel disease, Crohn's disease. Immune disorders include
XX rheumatoid arthritis, dermatitis, multiple sclerosis, neuritis,
XX orphalmia, systemic lupus erythematosus, insulin dependent diabetes
XX mellitus, and Guillain-Barre syndrome. The antibody is used to inhibit an
XX immune response. The polynucleotides can be used for production of the
XX protein, as diagnostic tools and to construct transgenic and knockout
XX animals and can also be used in chromosome mapping.
XX
XX Sequence 6454 BP; 2119 A; 1234 C; 1303 G; 1798 T; 0 other;

alignment_scores:
Quality: 2480.50

Length: 2266

Ratio: 1.964 Gaps: 58
Percent Similarity: 55.737 Percent Identity: 29.170
alignment_block:
US-09-737-246-2 x AAS08334 ..
Align seg 1/1 to: AAS08334 from: 1 to: 6454

53 GlnuValaAspProValaAspLeuGluAspTyrLeuIleThrHisProle 69
11 AAAGTGTGTGAGCCCTGGACATGAGATGTTATGTGCCAAAGAAAC 60
69 uAlaValaAspSerGlyProLeuArgAspLeuIleGluPheProProAspA 86
61 CCAATTTACAGGAGCCCTCCGAGATGTCGTATGTGTCCAAATGGAAG 110
86 sPllleGluValaValTyrSerProArgAspCysArgThrLeuValSerAla 102
111 ATATATCTATCTCGTGTAGTGTGTCAGCGCAGAACGGTCAGTCTACT 160
103 ValProGluGluSerGlu.....MetAspProHisValArgAspCy 116
161 GTACCAGAGAGATGCTGAAAGAGGCCCGCCAGAGTTATTGTTAAGAGTG 210
116 sIleArgSerTyrThrGluAspTyrAlaIleValIleArgLysTyrHisL 133
211 TATTAAGCTATATAGCAGATGTCGACGTGTAACTACAAATATAGAG 260
133 ySleuGlyThrGlyPheAsnProAsnThrLeuAspLysGlnLysGluArg 149
261 ACTTCTGTGGGAGACTTT.....CGAATGTTGCCATGTAATCTTTGAGA 304
150 GlnLysGlyLeuProLysGlnValPheGluSerAspGlnLalProAspG 166
305 CCAGAAAAGATGCTCAATCATGTAATTTGAGATGATGA..... 343
166 yAsnSerTyrGlnAspAspGlnAspAsp..... 175
344 ...GACTGTGAGAAAGATGAGGACATCTCTTATATGTTTCAGAGG 389
175 175
390 GTGGTGTGATMAAACAGGCTGTGCATTAAGCAATGTAAATAGTACC 439
176LeuLysArgArgSerMetSerLeuAspAs 185
440 ATCAACATTAACCATGAAGGATTTCAAGAGACGATATTATTACTTGACCCA 489
185 pThrProArgGlySerTyrPalaCysSerIlePhe...AspLeuLysAsn 201
490 ACTTCGTGAGCGTTCATATATTTCATATCTCTTAAGATGAGAAAAAT 539
201 erLeuPro.....AspAlaLeuLeuPro..... 208
540 CAAGAAGATGGAAGGTTGCATCTACTTGGAGCCGTCGATGTGTTGTT 589
209AsnLeuLeuAspArg 213
590 CAGTGGCCCAAAATGCCCGCATGCTTTTGAACSTCAAGATGTAGATTA 639
213 g..... 213
640 ATATAGSCATTAATCTGCTGCTGAAGTGAAGCAAGAAATGAGAGATGCT 689
213 213
690 TGATACTTTGAAAAAGATTAATTCAGATCAACACSCGACGTTAGTTCA 739
214 ThrGlnAsnGluGluIleAspArgGlnAspAspGlnArgLysSerAs 230
740 GAAAGAAAGAGAGCGTGAAGAACAGCAGAGATGGAATGAACTGACAGCA 789
230 pArgHisLysGluLeuPhe.....AlaLeuHisArgPro 241


```

790 AGGAAAAAGCCGAGAACATCATGCGAGTTTGGAAAGAGAGCATGCAATCCGG 839
      |||      :
241 er.....ProAspGluGluProIleGluArgLeuSer... 252
      :|||      |||
840 AACTGATGAGTATGGAAGAGAACTGAACAACATAACAACTCAGTGA 889
252 ..... 252
890 GGAGATGAGAGACAGAAATCTCTTTCTTTGATTCAGAAAGTTAGAGTT 939
253 .....ValProAspIlePro.....LysGluHisPheG 262
      |||      |||
940 GGACATTTTCAGGAATGGAACCTGATATAAGCCATTGAGAGAAATGCA 989
262 LysGluArgLeuLeuValLysCysLeuSerLeuLysPheGluIle..... 276
      :|||      :|||
990 ATAAACGTTTCTGCTGGTGAATGGCCATGATTACTTCAATATCTTGGGC 1039
277 .....GluIleGluProIlePhe 282
1040 CAATTTGGAGACATGCAAAAGAGCCACCAACAAATGTTGAGCCCTTTT 1089
282 eAlaSerLeuAlaLeuTyrAspValLysGluLysLysIleSerGluA 299
      :|||      :|||
1090 TATCATCTTGCCTTATTTGATGATTAAGACAAATGTAAAGATTTCACAG 1139
299 snPheTyrPheAspLeuAsnSerGluGlnMetLysGlyLeu..... 313
1140 ACTTTCAATGATAGACCTGAAATCCCATCTGTCCGTAATGCTGTGGGC 1189
314 .....ArgProHisValProPr 319
1190 TCTTCAACCCACATGCGCAGTAGCGATGAGCCAAAGGCTCTTCAACCGA 1239
319 oAlaAlaIleThrThrLeuAlaArgSer.....A 329
1240 ATCTTACATTCATGGAATTCGGAATTCAGTTACGTACATACAAACAG 1289
329 laIlePheSerIleThrTyrProSerGlnAspValPheLeuValIleLys 345
1290 GAATTTTTCAGATGACGAAATCCACATCCTGAATTTTCTAGTGGCCAGA 1339
346 LeuGluLysValLeuGlnGlnLysAspIleGlyLysAlaGluProTyr 362
1340 ATTGAAAGGTTACTA...CAGGGAACATTTACACACGTGGCAAGACCTTA 1386
362 rMetIlePheLysGluAlaAspAlaThrLysAsnLysGluLysLeuGlu 379
1387 TATTC.....AAAATTCGTGATCCAGTAAAGACGGCCAGAAAGTGCA 1430
379 yLeuLysSerGlnAlaAspGlnPheCysGlnArgLeuGlnLysTyrArg 395
1431 GG.....ACAGCTAAACAGAGTGTAGCCGCTTGGACAAATACAGA 1471
396 MetProPheAlaIleThrAlaIleHisLeuMetAsnIleValSerSerAl 412
1472 ATGCCCTTGGCTTGGGCGCCAGACCAATTTTCANA.....GAAATCCA 1515
412 asGlySerLeuGluArgAspSerThrGluValGluIleSerThrGlyGluA 429
1516 AGGCTCTCTTGATTCGAT..... 1534
429 rGlyGlySerTyrPheSerGluArgAsnSerSerIleValGlyArgArg 445
1535 .....GGGGA... 1540
446 SerLeuGluArgThrThrSerGlyAspAspAlaCysAsnLeuThrSerPh 462
1541 .....TT 1542
462 eArgProAlaThrLeuThrValThrAsnPheLysGlnGlnGlyAspA 479
      |||

```

```

1543 TTCTCT.....CTGTATAAACAGACAGTAGCA 1571
479 rLeuSerAspGluAspLeuTyrLysPheLeuAlaAspMetAlaArgPro 495
1572 AGCTTTAAAGTGAAGACATTTCTCAAGTTGCTCTCAGAAATATAAGAGCA 1621
496 SerSerValLeuArgArgLeuArgProIleThrAlaGlnLeuLysIleAs 512
1622 GAAAGAGC.....AACTGCAGATATATCTGGGAGCTAAACATCAC 1665
512 pIleSerProAlaProGluAsnProHisTyrCysLeuThrProGluLeu 529
1666 AGTAAATGTTCTGCTGGATTTATCAAAATGTATATCTCTCTCTATAG 1715
529 eugInValLysLeuTyrProAspSerArgValArgProThrArgGluIle 545
1716 TGCCCTTGAAGCCTTTGAAGAAATATGCCAAATATTTACTGTGGAGATT 1765
546 LeuGluPhe...ProAlaArgAspValTyr...ValProAsnThrTyr 560
1766 GAAGAGTTTGTCCAGAAATGACAAATATTTGATTCATTTACTATTTA 1815
560 rArgAsnLeuLeuTyrIleTyrProGlnSerLeuAsnPheAlaAsnArg 577
1816 CAAAACCATCTGATGTATATCCCGTCAATTAATAATGATAGACAG 1865
577 In.....GlySerAlaArgAsnIleThrValLysValGlnPheMet... 590
1866 AAACATTTGCCAAGCAGACAAACATTCAGCTGTGGGAAATCCGGGAT 1915
591 TyrGlyLysAspProSerAsnAlaMetProValIlePheGlyLysSer 607
1916 TCAGATGAAGATGACGTAGTCCCTAAAGTATATTATGAAAGACCTGC 1965
607 rCysSerGluPheSerIleGluAlaTyrThrAlaValValTyrHisAsn 624
1966 AGGCTGTGTTTAAACCAAAATGCTTATGCTGTGCTGCATCACACACC 2015
624 rGserProAspPheHisGluGluIleLysValLysLeuProAlaThrLeu 640
2016 AAAATCCAGATTTATGATGAGATTAAATTAAGCTTCCATTCACCTTA 2065
641 ThrAspHisHisLeuLeuPheThrPheTyrHisValSerCysGln... 656
2066 CATCAAAACATCATTTGCTTTCACATTTTATCAGTAGTGTGAAT 2115
657 .....GlnLysGlnAsnThrProLeuGluThrPro 667
2116 TAACACAAAGGGAACAACCAAAAGCAAGACACA...GTTGAACCTCCAG 2162
667 alGlyTyrThrThrIleProMetLeuGlnAsnGlyArgLeuLysThrGly 683
2163 TTGGGTTTGGCTGGTACCTTTGCTGAAAGATGTGAATCATCACATTTT 2212
684 GlnPheCysLeuProValSerLeuGluLysProProGlnAlaTyrSerVa 700
2213 GAGCAGACAGTGGCCAGTTTCCGCAATCTTCCCCAGAGCTACTGATCT 2262
700 lLeuSerProGluValProLeuPro.....GlyMetLysTrpValA 714
2263 GAATGATGACAGAAATCAAGAGCAATGTAACTGATATTAATTAATGGGTAG 2312
714 sPAsnHisLysGlyValPheAsnValGluValAlaValAlaValSerIle 730
2313 ATGGTGCAGAGCCTTTGTGAAGTTTAAAGCCACTTAGATCTTACCAT 2362
731 HisThrGlnAspProTyrLeuAspLysPhePheAlaLeuValAsnAlaLe 747
2363 TACACTCAGATGTCGATGTGCACAAATCTTC..... 2395
747 uAspGluHisLeuPheProValArgIleGlyAspMetArgIleMetGluA 764
2396 ....CATCATTTGCCAGCTGATTCAGTCAGGCTGCAAGAGATP..... 2434

```

```

764 snAsnLeuGluAsnGluLeuLysSerSerIleSerAlaLeuAsnSerSer 780
2435 .....CCAGGGAGCTCATTAATTAATTAAGGTTGGCATGCCATG 2476
781 GlnLeuGluProValValAlaArgPheLeuHisLeuLeuAspLysLeu1 797
2477 GAGATCCAGTCATGATACAGTTTCTACCTGTAATTTCTATGCCACTGTT 2526
797 GluLeuValIleArg.....ProProValIleAlaG 808
2527 CCGAGTTCTCACAAATATGACCCATGACATGACGTTCTATCAACTGCA 2576
808 LysIleValAlaAsnLeuGluGlnAlaSerPheGluAlaMetAlaSerIle 824
2577 CCATGGTT.....CTCTTAACATATTT 2596
825 IleAsnArgLeuHisLysAsnLeuGluGlnHisAspGlnHisGlyAla 841
2597 GTATCAAGTCCCAT.....GAAGAAGGCTT 2622
841 GAsnSerLeuLeuAlaSerTyrIleHisTyrValPheArgLeuProAsnT 858
2623 GGATAGTTTCTAAGATCATTCATTAAGTATAGCTTCCGA..... 2662
858 hTyrProAsnSerSerSerProGlyGlyLeuGlyLysSerVal 874
2662 ..... 2662
875 HisTyrAlaThrMetAlaArgSerAlaValAlaProAlaSerLeuAsnLe 891
2662 ..... 2662
891 uAsnArgSerArgSerLeuSerAsnSerAsnProAspIleSerGlyThrP 908
2662 ..... 2662
908 roThSerProAspArgPglValAlaArgSerIleIleGlySerLysGlyLeu 924
2663 .....CCTGAA..... 2668
925 AspArgSerAsnSerTyrValAsnThrGlyGlyProLysAlaAlaProT 941
2668 ..... 2668
941 pGlySerAsnProSerProSerAlaGluSerThrGlnAlaMetAspArgS 958
2669 .....AACCGAGTGT..... 2680
958 erCysAsnArgMetSerSerHisThrGluThrSerSerPheLeuGlnThr 974
2680 ..... 2680
975 LeuThrGlyArgLeuProThrLysLysLeuPheHisGluGlnLeuAlaLe 991
2681 .....CCTCAGGCCCACTGATACATGAACCCCTGGTAC 2715
991 uGlnTyrValValLysSerGlySerValArgGluSerAla..... 1004
2716 TACGATGATAGCAATA.....TTGAACAGTCTGCAGATTTTAT 2756
1005 .....LeuGlnGlnAlaTyrPhePhePheGluLeuMetVal 1016
2757 CATATAACAATTCCTAAGACTCATGTTTCTTTGAAATAATATGCA 2806
1017 LysSerMetValHisHisLeuTyrPheAsnAspLysLeuGlnAlaProAr 1033
2807 AAGTCAATGGCCACACTTGTGGAGAGATAAGATTAAAGCTTCCCG 2856
1033 GlySerArgPheProGluArgPheMetAspAspIleAlaAlaLeuValS 1050
2857 AGCCGAGAGATTTCCGAGACATATCATGCTTACATTCACGCTTC 2906

```

```

1050 erThrIleAlaSerAspIleValSerArgPheGlnLysAspThrGlnMet 1066
2907 TTGCAATAATTCCTCCCATGCTACTATTCGTATGGCGAGATCCCGATGAG 2956
1067 ValGluArgLeuAsnThrSerLeuAlaPhePheLeuAsnAspLeuLeuSe 1083
2957 TCCAGAAATGAGACTATAGTTTGGCTAGCTTCTCGAAGCGCTGTTTAC 3006
1083 rValMetAspArgGlyPheValPheSerLeuIleLysSerCysTyrLysG 1100
3007 ACTAATGATAGAGATTATTTTCAATTAAATAAAGACTATATATATC 3056
1100 LysValSerSerLysLeuTyrSerLeuProAsnProSerValIleValSer 1116
3057 GATTCAGCCCAAA.....GATCTAAGCTTGGCTGAA 3091
1117 LeuArgLeuAspPheLeuArgIleIleCysSerHisGluHisTyrValTh 1133
3092 TACAAGTTTGAAATTCTGCAAACAATTTGCATACACAGACATTAACATTC 3141
1133 rLeuAsnLeuProCysSerLeuLeuThrProProAlaSerProSerProS 1150
3142 TCTGAACCTTGCCAATGGCATTTCGAAACCTAAACTGCAG..... 3181
1150 erValSerSerAlaThrSerGlnSerSerGlyPheSerThrAsnValGln 1166
3182 .....CGGTTCAA 3190
1167 AspGlnLysIleAlaAsnMetPheGluLeuSerValProPheArgGlnG 1183
3191 GATTCAAATCTTGAA.....TACAGTTTATCAGATGATGATTGCAAGCA 3234
1183 nHisTyrLeuAlaGlyLeuValLeuThrGlnLeuAlaValIleLeuAsp 1200
3235 TCACCTCTTGCTGCTCTACTCTGAGCGAAACTCCACTGCTCTGACG 3284
1200 roAspAlaGluGlyLeuPheGlyLeuHisLysLysValIleAsnMetVal 1216
3285 ACAATTATGAC.....ATCAGATATACAGCTATCTCTGTATAT 3322
1217 HisAsnLeuLeuSerSerHisAspSerAspProArgTyrSerAspProG 1233
3323 AAGATCTTTTGATTAACAACATGCATTGGACACAAGATACGACACAAGAA 3372
1233 nIleLysAlaArgValAlaMetLeuTyrLeuProLeuIleGlyIleLeu 1250
3373 CCAACAAGCCAAATAGCACAAATTGACTCCCTTGTGGACTCTTT 3422
1250 etGluThrValProGlnLeuTyrAspPheThrGluThrHisAsnGlnArg 1266
3423 TGGAAATATATACAGCATTTAGCAGCTGAGATACCTGTATTTCTTGTCGA 3472
1267 GlyArgProIleCysIleAlaThrAspAspTyr..... 1277
3473 GCCATGCTATTTCTGCATCCAGAGATGAGTTTCCATGTGGCTTACTTC 3522
1277 ..... 1277
3523 ACCTGCCAATAGAGGAGTCTGAGCACTGACAAAGACCCGCTTATGGGT 3572
1278 .....GluSerGluSerGlySerMet 1284
3573 CTTTTCAAATGCGCATGGAATTAGAGAGATTCACAGAGCTTCCCTC 3622
1285 rLeSerGlnThrValAlaMetAlaIleAlaGlyThrSerValPro..... 1299
3623 ATCCAGAA.....GGAGCAACAGAGATTTCACAGATCA 3654
1300 .....GlnLeuThrArgProGlySerPheLeuLeuThrSerT 1312
3655 GGGCAACACTGGTGAAATAATACCCGACAGAGTCTT.....ACAAGA 3695
1312 hTyrSerGlyArgGlnHisThrThrPheSerAlaGluSerSerArgSerLeu 1328

```

3696 GINAGTATCCCGATTAACCGCGCTGGATCATGTAAATCAGAAAGCCCTC 3745
1329 LeuILECysLeuLeuThrValLeuLysAsnAlaAspGluThrValLeuGlu 1345
3746 CTGATGTGCTACTGTATATAGTAAAAATGATTTCTGGAAGATACCTCTCT 3795
1345 nIystrPheThrAspLeuSerValLeuGluLeuAsnArgLeuLeuAspL 1362
3796 AACTTACTGGAATAAGTATACCTACCTCAGAGCTCAATAACATTCTATATC 3845
1362 euLeuTyrlLeuCysValSerCysPheGluTyrlLysGlyLysValPhe 1378
3846 TTTTAAAGATATGCTGTTCCTTACCTTATGATATATGGGAAAAAACAATA 3895
1379 GluArgMetAsnSerLeuThrPheLysLys.....SerLysAspMetArg 1393
3896 GCAAGGGTCATGATGCTCGGCTGTCAAAAACATTCGGATATGACCGGAA 3945
1393 gAlaLysLeuGluGluAlaIleLeuGlySerIleGlyAlaArgGluGlu 1410
3946 ATCGCAAAACCATGCGCTCTTCGAAACAGATCAGAGATGATGACAGGCC 3995
1410 etValArgArgSerArgGlyGluLeuGluArgSerProSerGlySerAla 1426
3996 GCGTTCAG.....CACTTAGTAGCCTAGAAAAGTTCA 4027
1427 PheGlySerGlnGluAsnLeuArgTrpArgLysAspMetThrHisTrpArg 1443
4028 TTT.....ACACTTAATCACAAGTTC 4047
1443 gGluAsnThrGluLysLeuAspLysSerArgAlaGluIleGluHisGlu 1460
4048 TACAACAACATAA.....GCAGACATTTCCACCAAG 4079
1460 IeLeuLysPheGlyAsnLeuAlaThrGluAlaAsnLeuIleLeuAsp 1476
4080 CACTTCTTAAGGCATATACGACTGTAAGTTCCCTTACAGTACTAGAC 4129
1477 ThrLeuGluIleValAlaGlnThrValSerVal.....ThrGlu 1489
4130 ACCATATCATTTTTCACCTACGTCCTTCAAGACCAACTTTAAATATATA 4179
1489 uSerLysGluSerIleLeuGlyGlyValLeuLysValLeuLeuHisSerM 1506
4180 TGCCCATTAACCCCATTAATGAAAAAGTGTGATATACATCTTGCTTTC 4229
1506 etLacLysAsnGlnSerAlaValTyrlLeuGlnHisCysPheAlaThrGln 1522
4230 TTTAAAAATGACATCTGAAGTGCCTGTAACATGATATTTGCTCACATG 4279
1523 ArgAlaLeuValSerLysPheProGluLeuLeuPheGluGluGluThrGlu 1539
4280 AGAGCTTTCATCAGTAAGTTCTCTTACGATTTTTCAAAGAGAGATATA 4329
1539 uGlnCysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSerI 1556
4330 CATGTGTGCTGATTTTCTATGAGGTTTAAAGTCTGCATCATGCAAGA 4379
1556 IeGlyThrIleArgSerHisProSerAlaSerLeuTyrlLeuLeuMetArg 1572
4380 TTAGCTCAACAGGAATGAAAGCATCTGCACTTTGTATCTTTGTATGAGA 4429
1573 GlnAsnPheGluIle.....GlnAsnAsnPheAlaArgValLysMetGlu 1587
4430 AACAACCTTTGAGTATACCAAAAGAAACCTTTTGAGSACACATCTATA 4479
1587 nValPrometSerLeuSerSerLeuValGlyThrSerGlnAsnPheAsnG 1604
4480 GATAAATAATTTGGTGTAAAGCAACTGATAGTATGACACTAAGCGAG 4529
1604 IuGluPheLeuArgSerLeuLysThrIleLeuThrTyrlAlaGluGlu 1620

4530 GATCAAGATTTTCAGAGCTTTTATTCATTTATCAATTAATTTGCAATAGT 4579
1621 AspLeuGluLeuArgGluThrThrPheProAspGlnValGlnAspLeuVal 1637
4580 CACAGACCTATGAAGGCACTGCTTTCCTCCGACAGAAAGTCAAGACTTGAC 4629
1637 I PheAsnLeuHisMetIleLeuSerAspThrValLysMetLysGluHisG 1654
4630 CAAGAGATCCGACACTGTCTTATGGCCACTGCCCAATGAAAGAGCATG 4679
1654 IuGluAspProGluMetLeuLeuLeuAspLeuMetTyrlArgIleAlaLysGly 1670
4680 AGAAAGACCTGAAATGCTATATGATCTCCAGTATGACTTACGCCAAGTCC 4729
1671 TyrGlnThrSerProGlu.....ArgLeuThrTyrlLeuGlnAsnMetAlaGlu 1686
4730 TATGCAAGACCCCAAGACCTCAGGAAAAACCTGCTTATGATGATGCCAA 4779
1686 yLysHisSerGluArgSerAsnHisAlaGluAlaGluCysLeuValH 1703
4780 GATTCATGTAATAAATGAGATTTTTCAGAGCGTGCATGTGTATGTCC 4829
1703 I sSerAlaIleValAlaIleGluTyrlLeuSerMetLeuGluAspArgLys 1719
4830 ATGTAAGACGCTCTAGTTGCAGAGTTT.....CTTCATCGAAAAAA 4870
1720 TyrLeuProValGlyCysValThrPheGlnAsnIleSerSerAsnValIle 1736
4871 TTATTTCTTAAGGATGTTCAGCGCTTCAGAAAAATTCTCCCAATATAGA 4920
1736 uGluGluSerAlaValSerAspAspValValSerProAspGluGluGly 1753
4921 TGAAGAAAGCAATGAAGAAAGATGCTGGATGATGAT..... 4960
1753 IeCysSerGlyLysTyrlPheThrGluSerGlyLeuValGlyLeuGlu 1769
4961GTCCATTATAGTGAAGAGTCTTGCGGAGTCTGTGAA 4999
1770 GlnAlaAlaIleAspPheSerMetAlaGlyMetTyrlGluAlaValAsnGlu 1786
5000 CAATGTGTGATGCTTATGSAAGCGACAGCTTATGAAATATTTCTGCA 5049
1786 uValTyrlLysValLeuIleProIleHisGluAlaAsnArgAspAlaLysL 1803
5050 GATTCCCAAGTGAATGCTTCAATTTATGAAACGTCGAGATTGAGA 5099
1803 yLysSerThrIleHisGlyLysLeuGlnGluAlaPheSerLysIleVal 1819
5100 AACTTACTCAAGTTTATGAACTCTTATGAGCTTACACAAAAATCTGT 5149
1820 HisGlnSerThrGlyTrpGluArgMetPheGlyThrTyrlPheArgValGlu 1836
5150 GAAGTTATGCATACAAAAAGACACTTTTACGCACTTCTTCACAGTTGC 5199
1836 yPheTyrlGlyThrLys...PheGlyAspLeuAspGluGluPheValT 1852
5200 CTTTATATGCCAATCTTTTGTGAAGAAAGATGGAAGAGATACATCT 5249
1852 yTyrlGluProAlaIleThrLysLeuAlaGluIleSerHisArgLeuGlu 1868
5250 ATAAAGAACCAAAAGCTCAGTCCCTCAGAAATTTCTTGAGCACTGT 5299
1869 GlyPheTyrlGlyGluArgPheGlyGluAspValAlaGluValIleLysAs 1885
5300 AAACCTTATGTGAAAAAGTTGTTACGAGATGTCAAAATATTTCTGAGA 5349
1885 pSerAsnProValAspLysCysLysLeuAspProAsnLysAlaTyrlIleG 1902
5350 TTCAGACAAAGTAAATGCAAGAGCTTGATCCAAATATGCTCATATATC 5399
1902 IuIleThrTyrlValGluProTyrlPheAspThrTyrlGluMetLysAspArg 1918
5400 AAGTTACTTATGTGAAGCTTACTTGTATGACAAAGAACTCACAGAAAG 5449

[illegible][illegible]

```

546 LeuGluPhe...ProAlaArgAspValTyrVal...ProAsnThrThrTyr 560
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1909 GAGCAATTGTGGCTGCATACCAAAACACACACCTTACACCATCTA 1958
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
560 rArgAsnLeuLeuTyrIleTyrProGlnSerLeuAsnPheAlaAsnArg 577
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1959 CACCAATCACCTTACGTTTATCTTACGTAAGTGAATGACGACATCAGA 2008
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
577 In.....GlySerAlaArgAsnIleThrValLysValGlnPheMet... 590
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2009 AGCTTTTGGCAAGGCTACAAATATTGCATTTGCATTGAATTCAAAGAT 2058
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
591 TyrGlyLysProSerAsnAlaMetProValIlePheGlyLysSer 607
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2059 TCAGATGAGGAAGACTCTCCAGCCCTTAAGTGCAATTTATGGCAGACCTGG 2108
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
607 rCysSerGluPheSerLysGluAlaTyrThrAlaValTyrHisAsn 624
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2109 TGGGCCAGTTTTCACAAAGAGCGCTTGTGTCAGTTTTCACCAATCAC 2158
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
624 rSerProAspPheHisGluGluIleLeuValLysLeuProAlaThrLeu 640
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2159 AAACCCCAAAATTTTATGATGAGATTAAATAGAGTTGCCACTCAGCTG 2208
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
641 ThrAspHisHisLeuLeuPheThrPheTyrHisValSerCysGln 657
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2209 CATGAAGACACACCTGTTGCTCATCTCTCCATGTCAGCTGGACAA 2258
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
657 n.....LysGlnAsnThrProLeuGluThrProValG 668
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2259 CTCAGTAAGAGAACGACGAAAGAGGAGTGTCTTGAAACCAAGTTG 2308
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
668 LTTTThrThrIleProMetLeuGlnAsnGlyArgLeuLysThrGln 684
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2309 GCTACTCTGGCTCCCTCCCTCTGAAAGACGAGAGGCTGTCACAGCGAG 2358
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
685 PheCysLeuProValSerLeuGluLysProPro.....GlnAlaTyrSe 699
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2359 CAGCAGATCCCGGTCTCGGCAACCTCTCTCGGCTATCTTGCTGCTACCA 2408
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
699 rValLeu.....SerProGluValProLeuProGlyM 710
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2409 GGAGCTTGGATGGCAGGCAATTATGCTCGGAAAT..... 2445
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
710 eLysTrpValAspAsnHisLysGlyValPheAsnValGluValAla 726
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2446 ..AAATGGGTAGATGAGGCAAGCCACTGCTGAATAATTCCACTATCTG 2493
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
727 ValSerSerIleHisThrGlnAspProTyrLeuAspLysPhePheAla 743
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2494 GTTTCACAGTGTATACCTCAGAGTACAGCATTTACATATTTTTCACATA 2543
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
743 uValAsnAlaLeuAspGlnHisLeuPheProValArgIleGlyAspMet 760
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2544 CTGTGAGAAACCGAATCTGGA..... 2565
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
760 rGluMetCysAsnLeuGluAsnGluLeuLysSerSerIleSerAla 776
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2566 .....GCCCAAGCCTTAGAAAGAACTGTAAAGTACCTTAAGAGT 2607
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
777 LeuAsnSerSerGlnLeuGluProValArgPheLeuHisLeuLeu 793
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2608 CTGCATGCGATGGAAAGCCACGTGATGATGCGCTTCCGCCACATCTCT 2657
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
793 uAspLysLeuIleLeuLeuValIleArgProProValIleAlaGlyGln 810
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2658 AAACCAAGCTGTCCGAGTCTCACACAGA..... 2685
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
810 LeuValAsnLeuGlyGlnAlaSerPheGluAlaMetAlaSerIleLeuAsn 826
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2686 .....GCCACACAGGAAGAGTCGCGTTAAACCTGACT 2718
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
827 ArgLeuHisLysAsnLeuGluGlyAsnHisAspGlnHisGlyArgAsnSe 843

```

```

2719 CGGGTCATTATTCATGTGGTGGCCAGTCCCATAGAGAGATTGGAG 2768
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
843 rLeuLeuAlaSerTyrIleHisTyrValPheArgLeuProAsnThrTyr 860
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2769 CCAGTTGAGTCATATGTGTAAAGTACGGCTAT..... 2799
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
860 rAsnSerSerSerProGlyProGlyLeuGlyGlySerValHisTyr 876
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2799 ..... 2799
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
877 AlaThrMetAlaArgSerAlaValArgProAlaSerLeuAsnLeuAsnArg 893
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2799 ..... 2799
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
893 gSerArgSerLeuSerAsnSerAsnProAspIleSerGlyThrProHis 910
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2799 ..... 2799
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
910 eProAspAspGluValArgSerIleIleGlySerLysGlyLeuAspArg 926
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2799 ..... 2799
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
927 SerAsnSerTrpValAsnThrGlyProLysAlaAlaProTrpGlySe 943
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2800 .....AAGCTCAGCCATATGTCTG 2819
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
943 rAsnProSerProSerAla.....GluSerThrGlnAlaMetAspArg 958
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2820 CTCTGAATACAAAGACAGTGCATGAAGAAGTACCACCAATTCATG..... 2862
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
958 eCysAsnArgMetSerSerHisThrGluThrSerSerPheLeuGlnThr 974
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2863 .....ACCACGATTCACAGCCT 2880
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
975 LeuThrGlyArgLeuProThrLysLysLeuPheHisGluLeuAla 991
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2881 TCTGCCGATTTCTCACCACGACCAAACTA..... 2910
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
991 uGlnTrpValValCysSerGlySerValArgGluSerAlaLeuGlnGln 1008
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2911 .....CTGACAGTACT 2920
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1008 LatrPhePhePheGluLeuMetValLysSerMetValHisHisLeuTyr 1024
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2921 CATGTTTCTTGTGATGTACTGATCAATCTATGCTCAGCATTTGATA 2970
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1025 PheAsnAspLysLeuGluAlaProArgLysSerArgPheProGluArg 1041
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
2971 GAGAACTCCAAAGTTAAGTTGCTGCGAAACGACAGATTTCTGCACTCTA 3020
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1041 eMetAspAspIleAlaLeuValSerThrIleAlaSerAspIleVal 3058
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
3021 TCATCATGACAGTGGAAACCGTGTAAATATGCTGATGCCACACATGACCT 3070
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1058 eArgArgPheGlnLysAspThrGluMetValGluArgLeuAsnThrSerLeu 1074
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
3071 AGAAGTTTCAGATTAATCCAGAGGCATCTAAGAACGGAATTCATACCTT 3120
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1075 AlaPhePheLeuAsnAspLeuLeuSerValMetAspArgGlyPheVal 1091
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
3121 GCTGTCTTCAACAAGATGTTTACCTTCATGTGACAGAGGGCTTTGTCTT 3170
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1091 eSerLeuIleLysSerCysTyrLysGlnValSerSerLysLeu.....T 1106
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
3171 C.....AAGCAGATCACACACTACATTAAGCTGTT 3199
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1106 ySerLeuProAsnProSerValLeuValSerLeuArgLeuAspPheLeu 1122
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
3200 TTGCTCTTGAGAGACCCAAAGACCTCTTGAATACAGTTTGAATTTCTC 3249
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1123 ArgIleIleCysSerHisGluHisTyrValThrLeuAsnLeuPro-Cys 1139
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```
3250 CGTGAAGTGCACCATGACATTAATTCCTGGAACCTTACCAATGCC 3299
1139 erleuethrProProAlaserProSerProSerValSerSerAlatr 1155
3300 ATTTG..... 3304
1156 SerGlnSerSerGlyPheSerThrAsnValGlnAspGlnLysIleAlas 1172
3305 ...GAAGAGCAGATTC.....AAGATACCAAGACCTCCAG 3339
1172 nmet..PheGlnLeuSerValProPheArgGlnGlnHisTyrLeuAla1 1188
3340 CTGGACTTACATTAACAGATGAGTCTGCAGAAACCACTTCTGGTGG 3389
1188 yLeuValLeuThrGlnLeuAlaValIleLeuAspProAspAlaGluGlyL 1205
3390 ACCTGTTACTGAGGAGGTGGGACACCTCCAGGAGTTCCGGAG... 3435
1205 euhPheGlyLeuHisLysLysValIleAsnMetValHisAsnLeuSer 1221
3436 .....GTCCGTGATGCCATGACGTGCTCAAGAACCTGGTGATA 3477
1222 SerHisAspSerAspProArgTyrSerAspProGlnIleLysAlaArgVa 1238
3478 AACCATTTCTTTGATGACAGATAGCTTCACAGAGCCATCAGGCAGAGAT 3527
1238 lAlaMetLeuTyrLeuProLeuIleGlyIleIleMetGluThrValProg 1255
3528 AGCCACCTCTACTGCTCTGTTGGTCTGCTGATGAAAACGTCCAGC 3577
1255 lLeu.....TyrAspPheThrGluThrHisAsnGlnArgLysArgPro 1269
3578 GATCAATGTGAGGAGTGTGCACCTCCCTGTGAACGGGGCATGACT 3627
1270 lle.....CysIleAlaThrAs 1275
3628 GTGACAGGATGAATCCCTGGCTCTACACCTGTGAATCCCGCTGTGACGCC 3677
1275 pAspTyrGlnSerGlnUserGlySerMetIleSerGlnThrValAlaMetA 1292
3678 GCAGAGGAGGAGACCCCTGCAGACAGCCTGCAGCAAGACCTGCTGGGG 3727
1292 lAlaLeuAlaGlyThrSerValProGlnLeuThrArgPro..... 1304
3728 CCATCTCCGGCATGTCTCTCATATACAACTCAACCTCAAAACATCAAC 3777
1305 .....GlySerPheLeuLeuThrSerThse 1313
3778 AGTGTGAAGAAATGCTGATTCGAGAGGATCT...CTCATTAAGCAGATTC 3824
1313 rGlyArgGlnHisThrThrPheSerAlaGlnUser..... 1325
3825 GGGTACACACCTTCCGAAAGAAATAGTAGAAGAACATTCCTCGGATA 3874
1325 ..... 1325
3875 AGCACCACAAAGTAGACATTTGGAAATTCGGTGGTTCGTGTATAAA 3924
1326 .....ArgSerLeuLeuIleCysLeuLeuThrValle 1336
3925 CTTGACCACTGTGAGATTAAAGCCTTAAGATGATGTTCCTCTACATCTT 3974
1336 ulysAsnAlaAspGluThrValLeuGlnIlyStrPheThrAspLeuSerV 1353
3975 AAGAGACAGTCTGATGATGCTTTGTTTACATITGGAACAAGGCTTCA 4024
1353 alLeuGlnLeuAsnArgLeuLeuAspLeuLeuTyrLeuCysValSerCys 1369
4025 CATCTGAACCTTAATGATTTTATCAATATCTGAAGTCTGCTGCACCG 4074
1370 PheGluTyrLysGlyLysLysValPheGluArgMetAsnSerLeuThrPh 1386
4075 TTCAGTACATGGGGAAGCGATACATAGCCAGAAACAGGAATG..... 4116
1386 eLysLysSerLysAspMetArgAlaLysLeuGluGluAlaIleLeuGlyS 1403
4117 .....ATGCATGCCAAGATTGCAGAC.....CTGGGCA 4144
1403 erIleGlyAlaArgGlnGluMetValArgArgSerArgGlyGlnLeuGlu 1419
4145 GCCTG..... 4149
1420 ArgSerProSerGlySerAlaPheGlySerGlnGlnAsnLeuArgTyrPar 1436
4150 .....GATAACTCTCTCACTTTTAA 4169
1436 gLysAspMetThrHisTyrArgGlnAsnThrGluLysLeuAspLysSerA 1453
4170 CCACGGCTATGGCCAC.....TCGG 4189
1453 rGlaGluIleGlnHisGlnValLeuIleAspGlyAsnLeuAlaThrGlu 1469
4190 ACAGAGATGTTCTGCACACGATTAATCTTGAAGCCACATTTGCTACTGAG 4239
1470 AlaAsnLeuIleIleLeuAspThrLeuGluIleValVal..... 1482
4240 GTTGGCTGACAGCTCTGCACACGCTTCTCATATTACATGGCGTTTAA 4289
1483 .GlnThrValSerValThrGlnSerLysGlnSerIleLeuGlyLysValL 1499
4290 CAACACGCTCTGGCCCGACCATGGACATTAATCCTCTATGAAAAAAGTTT 4339
1499 eLysValLeuLeuHisSerMetAlaCysAsnGlnSerAlaValTyrLeu 1515
4340 TTGATGTCTACTCTGTGTTTCTTCAAAAACATAGTGTGAACGCCGTTTA 4389
1516 GlnHisCysPheAlaThrGlnArgAlaLeuValSerLysPheProGluLe 1532
4390 AAAATGTCTTCACTGCTTAAGCTCTTAATTATTAAGTTTCCCTCAAC 4439
1532 uLeuPheGluGluGluThrGlnGlnCysAlaAspLeuSerLeuArgLeuL 1549
4440 ATTCTATTAAGGAGAGGAGCGACATGTGTGCGCTCTGTGTACAGATTC 4489
1549 euArgHisCysSerSerSerIleGlyThrIleArgSerHisProSerAla 1565
4490 TCAAGTCTTAACCTCAAGCTGAGTGCATCAGAGGAGGCCCTCCAG 4539
1566 SerLeuTyrLeuLeuMetArgGlnAspPheGluLe.....GlyAsnAs 1580
4540 CTGCTCTACTCTGATGAGGAGAACACTTGTATTAACCTGGAAGAAGTCC 4589
1580 nPheAlaArgValLysMetGlnValPrometSerLeuSerSerLeuValG 1597
4590 CTTGTCTCGGACACATTTGCAGATCATATATCTGTGACCGACGTGTATG 4639
1597 lYthrSerGlnAsnPheAsnGluGluPheLeuArgArgSerLeuLysThr 1613
4640 CAGACGTTGTTGGCATTTGGGGGAACAGATTCACGACGTCCTGTATCC 4689
1614 lLeuLeuThrTyrAlaGluGluAspLeuGluLeuArgGluThrThrPhePr 1630
4690 ATCAACAACCTGTGCCACAGTACAGCGGCTTATTAAGCACACACACTTCTC 4739
1630 oAspGlnValGlnAspLeuValPheAsnLeuHisMetIleLeuSerAspT 1647
4740 CTCGTATGTGAGGACATTAAACCAAAAGATAGCAGCGTGTCTAATGGCCA 4789
1647 hrValLysMetLysGlnHisGlnGluAspProGlnIleLeuLeuAspLeu 1663
4790 CCGCCCAAGATGAGAGACATGAGAGCAAGCAAGAGATGTGTGGAGACTC 4839
1664 MetTyrArgIleAlaLysGlyTyrGlnThrSerProGlu...ArgLeuThr 1679
4840 CAGTACACCTTGGCCAAATCTATGACGACAGCCGCGAGCTCAGGAAGAC 4889
```


PR 11-APR-2000; 2000US-0196527.
PR 11-APR-2000; 2000US-0196528.
PR 11-APR-2000; 2000US-0547276.
PR 13-OCT-2000; 2000US-0240503.
PR 13-OCT-2000; 2000US-0240508.
PR 13-OCT-2000; 2000US-0240543.
PR 13-OCT-2000; 2000US-0240539.
XX
XX (ARBO-) ARBOR VITA CORP.
PI Lu P, Garmen JD, Candia AF.
XX WPI; 2001-367865/38.
DR P-PSDB; AAU04025.
XX
PT CLASP-5 polynucleotides, proteins and antibodies are used to prevent or
PT treat a CLASP-5 mediated disease, such as an autoimmune disease e.g.
PT Rheumatoid arthritis -
XX
PS Disclosure; Fig 9A; 188pp; English.
XX
CC The sequence encodes a protein presented as AAU04025, included in the
CC specification which relates to a cadherin-like asymmetry protein,
CC CLASP-5, which is a transmembrane protein of the immune system involved
CC in the formation of the immune gateway. CLASP-5, polynucleotides encoding
CC it and an anti-CLASP-5 antibody are used to prevent or treat a CLASP-5
CC mediated disease, such as an autoimmune disease caused or exacerbated by
CC increased activity of TH1 cells. These diseases may include Addison's
CC disease, rheumatoid arthritis, dermatitis, Multiple Sclerosis, Systemic
CC Lupus Erythematosus and Autoimmune Thyroiditis, Inflammatory conditions
CC (e.g Ischaemia-reperfusion), AIDS, syphilis, typhoid, Burkitt's lymphoma,
CC Immunodeficiency syndrome, AIDS, syphilis, typhoid, Burkitt's lymphoma,
CC Grave's disease and Insulin dependent diabetes mellitus. CLASP-5 is used
CC for inhibiting an immune response in a cell (T cell or B cell) by
CC interfering with the expression of a CLASP-5 gene in the cell, the
CC ability of a CLASP-5 protein to bind to another cell or the ability of a
CC CLASP-5 protein to bind to another protein. The antibody binding CLASP-5
CC is used to inhibit an immune response is a subject. The polynucleotide is
CC used to detect CLASP-5 expression in cells and for diagnosis of
CC diseases and disorders associated with aberrant expression of CLASP-5.
CC Note: The present sequence is included in the specification but is
CC not mentioned anywhere in the specification.
XX
SQ Sequence 2014 BP: 565 A; 469 C; 514 G; 466 T; 0 other;

alignment_scores:
Quality: 2245.50 Length: 619
Ratio: 4.098 Gaps: 2
Percent Similarity: 88.530 Percent Identity: 69.628

alignment_block:
US-09-737-246-2 x AAS07381 ..

Align seg 1/1 to: AAS07381 from: 1 to: 2014

1469 GIUALAAsnIleuIleIleuAspThrIleuGIuIleValIaGIuThrVa 1485
::: :::::::::::::::::::: ||| ::::::::::::::::::::
8 CAACGAGATTTAATCATCTCGATATGCAGGAAACATTATCCAGCGCAG 57

1485 lSerValThrIuSerIySGuSerIleIleuGIyGIyValIleuIyVal 1502
||||: ::| ::::::::::::::::::::| ::::::::::::::::::::|
58 CTCGGCTCGACGTGAAAGACAGCCTGCTGGAGGTGTCTGAGAGGTGC 107
CTCCTGAGTCTCTGAACGTGTATCAGATACCACTACCTGACTACTGC 157
108 TGGTGAATCTCTGAACGTGTATCAGATACCACTACCTGACTACTGC 157
1502 euIeuHisSerMetAlaCysAsnGIuSerAlaValIyIleuGIuHisCys 1518
||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
108 TGGTGAATCTCTGAACGTGTATCAGATACCACTACCTGACTACTGC 157
1519 PheAlaThrGIuArgAlaIleuValSerIyPheProGIuIleuPheGI 1535
|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
158 TTTCGACACGCTCCGTCCTCATTCGCCAAGTTGGAGAGACTTACTTTGCA 207
1535 uGIuGIuThrGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 1552

|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
208 AGAGAGGTGGAACAGTGTTCGACCTATGTCAACCAAGTCCTGCACACT 257
1552 ySerSerSerIleGIyThrIleArgSerHisProSerIaSerIeuIyTr 1568
|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
258 GCAGCAGCAGCAGCAGTGTACCCGAGCCCAAGCCTGTGCCACCTTTAC 307
1569 LeuIeuMetArgIuIleAsnPheGIuIleGIyAsnAsnPheAlaArgVal 1585
|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
308 CTCCTCATGAGTTCAGTTTGGAGCCACCAAGTAATTTTGCAAGATAA 357
1585 sMetGIuValProMetSerIeuSerIeuValGIyThrSerGIuAsn 1602
|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
358 GATGCAAGTAACCATGTCCCTGCAATCTTTGGTGGGAAGACACAGACT 407
1602 heAsnGIuGIuPheIeuArgArgSerIeuIyThrIleIeuPheIyAla 1618
|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
408 TTAAATGAAGACACCTGAGAGATCCTTGAGGACAATTTTGGCTTTCA 457
1619 GIuGIuAspIeuGIuIleuArgGIuThrThrPheProAspGIuValGIuAs 1635
|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
458 GAAGAGGACACAGCATGCGATGACTCTTTCCACCCAGCGTGGAGGA 507
1635 PheuValPheAsnIleuHisMetIleuSerAspThrValIyMetIySG 1652
|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
508 ACTCTCTGTAACTGTAATGCAATGATATGACACAGTGAATGAGGG 557
1652 IuHisGIuGIuAspProGIuMetIleuIleAspIeuMetIyArgIleAla 1668
|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
558 AATTTCCGGAAGATCCGAGATCTTATGATGATCTCATGACAAATGGC 607
1669 LysGIyThrGIuThrSerProGIu...ArgIeuThrTrPheuGIuAsnMe 1684
|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
608 AAGGTATACCAAGCATCTGATCTGCGGCTGACCTGGTCCAGAAAT 657
1684 tAlaGIyLysHisSerGIuArgSerAsnHisAlaGIuAlaIleGIuCys 1701
|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
658 GCGAGAAACACACACACMAAGAGTGTACACGAGGCTGCATGTGCC 707
1701 euValHisSerAlaIleuValAlaGIuThrIeuSerMetIeuGIuAsp 1717
|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
708 TGTGTGACGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 757
1718 ArgIyThrIeuProValGIyCysValThrPheGIuAsnIleSerSerAs 1734
|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
758 CACAGCTACCTGCCCGTGGGAGGTGAGCTTCCAGATATTTCTTCCAA 807
1734 nValIeuGIuGIuSerAlaValSerAspAspValIySerProAspGIu 1751
|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
808 TGTGTGTGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 857
1751 IuGIyIleCysSerGIyLysThrPheThrIuSerGIyIeuValGIyIleu 1767
|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
858 ATGGGTGTGCGGAGCGCAGTACTTCAACGAGAGTGGCTGTGAGGCTTC 907
1768 LeuGIuGIuAlaIleuValSerPheSerMetAlaGIyMetIyGIuAla 1784
|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
908 CTGAGACGAGCGCGGAGCTTTCAGACGAGCGCTTATATATGACAGT 957
1784 lAsnGIuValIyTrLysValIleuIleProIleHisGIuAlaAsnArgAsp 1801
|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
958 TTAATGAGGTCTACAAAGTGTATCCCATCCATCTTAAGAGCCATCGAAT 1007
1801 lAluIyLysIeuSerThrIleHisGIyLysIeuGIuIleuAlaPheSerLys 1817
|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1008 TCCGGAAGCTGACATCTACACAGCAAGCTGCAGAGAGCTTTCGAGAGC 1057
1818 lIleValHisGIuSerThrGIyTrpGIuArgMetPheGIyThrTrPhear 1834
|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1058 ATGCTTAAACAAGAT.....CATAAAGAAATGTTTGAACCTACTTCCG 1101
1834 gValGIyPheTrGIyThrLysPheGIyAspIeuAspGIuGIuIleuPhe 1851
|||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

1102 AGTTGGTTCTTGGATCCAAATTTGGGGATTGGATGACAGAGCTTG 1151
1851 ATTTTLYSGLUPROAIATLTHRLYSLLEUALAGLUILESERHISGREU 1867
|||||
1152 TCTACAAAGAGCTGCATATTCACCAAGCTTCTGCAATCTCACATGACTA 1201
1868 GUGLUGLPHETRYGLYLUAARGPHEGLYGLSUAPVALAGLUALILELY 1884
|||||
1202 GAGGCAATTTATGTCATAGTTTGGTGCAGAAATTTGGGAAGTCATTA 1251
1884 SASPERSANPROVALASPLYSCYLSYLEUASPPROASNLYSALATYRI 1901
|||||
1252 ACATCCACTCTCTGTGACAAAACCAAGTTGATCTTAACAAGGCTACA 1301
1901 LGLIILHETRYVALGILUPROTGYRPHASPTHTYRGLUMELLYASP 1917
|||||
1302 TACAGATCACTTTGTGAGCCCTACTTGTGAGTATGAGATGAAACAC 1351
1918 ARGILETHTRYRPHASPLYASNTRYASNLEUARARGPHEMETYRCY 1934
|||||
1352 AGGTCACATACTTTGAGAGAAATTTCAACCTCCGAGGTCATGATACAC 1401
1934 STHYRPHETHTHLEUASPGLYARGALAHISGLYGLULENHISGLUINP 1951
|||||
1402 CACCCCGTTCACCCCTGAGGCGGCGCTCGGGGAGAGCTGCATGAGCAGT 1451
1951 HELYSAIRGLYSTRHLEUTHRTHRSERHISALAPHEPROTYRILEYLS 1967
|||||
1452 ACAGAGACACACAGTCTTGACCTACTATGACAGGCTTCCCTACTACAG 1501
1968 THARVALASNAVALTHRLHISLYSGLULILELEUTHRPROILEGI 1984
|||||
1502 ACCAGAGATCAGCGTCATCCAGAGAGAGAGTGTGTTTGACACCGATGA 1551
1984 UVALAIALIEGLUASPMETGLNLYSLYSTHGLINGLULEUALAPHEALAT 2001
|||||
1552 AGTTCSCATTTGAAGACATGAAGAACACCCCTGCAGTTACAGTTGCCA 1601
2001 HRLHISGLASPROALASPROLYSMETLEUGLIMETVALLEUGLNGLY 2017
|||||
1602 TTAACCCAGAGCCGCTCATGCAAAAGATGCTTCAGATGCTGTCAGAGC 1651
2018 SERVALGLYTHRTHVALASGLINGLYPROLEUGLULVALAIALINAPH 2034
|||||
1652 TCTGSGGAGCTACTGTAATCAGGAGACCTGGAAGTACCCAAAGTTT 1701
2034 ELEUSERLUILEPROSERASPROLYSLEUPHEARGHISHSANLYSL 2051
|||||
1702 TTTGGCTGAAATTCCTGCTGATCCAAAACCTATCGACATCAACAAGT 1751
2051 EUARGLEUCYSPHELYSAPPHETHTHLYSARGCYSGLUASPALALEUARG 2067
|||||
1752 TGAGGTTTCTCTTAAAGAAATTCATCATGACATGTGGTGAAGCTGTAG 1801
2068 LYSANLYSSERLIEGLYPROVALGILNLYSGLUTYRGLINARGLULE 2084
|||||
1802 AAAAACAAGCGTCTCATCAGCGACAGACCAAGGGAATATCATCAGGAAC 1851
2084 UGLYLYS 2086
|||||
1852 CAAAAAG 1858

```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AACT74524

seq_documentation_block:

ID AACT74524 standard; cDNA; 6816 BP.

XX AAC74524;

XX 08-FEB-2001 (first entry)

DE Human ORFX ORF79 polynucleotide sequence SEQ ID NO:157.

KW Human: open reading frame; ORFX: detection; cytosstatic; hepatotropic;
KW vulnerability; antipsoriatic; antiparkinsonian; neurotropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antinephritic;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; burns; wounds; bone and cartilage damage;
KW bone damage; cartilage damage; nocturnal haemoglobinuria; burn; wound;
KW thrombosis; contraceptive; ss.

XX Homo sapiens.
XX OS
XX WO200058473-A2.
XX PN
XX 05-OCT-2000.
XX PD
XX 31-MAR-2000; 2000WO-US08621.
XX PF
XX 31-MAR-1999; 99US-0127607.
XX PR 02-APR-1999; 99US-0127636.
XX PR 05-APR-1999; 99US-0127728.
XX PR 30-MAR-2000; 2000US-0540763.
XX PA (CUBA-) CURAGEN CORP.
XX PI Shinkets RA, Leach M;
XX DR WPI: 2000-602362/57.
XX DR P-PSDB: AAB40315.
XX DX
XX PT Novel nucleic acids and peptides derived from open reading frame X,
XX PT useful for treating e.g. cancers, proliferative disorders,
XX PT neurodegenerative disorders and cardiovascular disease -
XX PS Claim 5; Page 502-506; 5507pp; English.

CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytosstatic; hepatotropic; vulnerability;
CC antipsoriatic; antiparkinsonian; neurotropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antinephritic; antibacterial; antifungal; antirheumatic; antithyroid;
CC antianemic; gene therapy; cancer; proliferative disorder; hypertension;
CC neurodegenerative disorder; osteoarthritis; graft vs host disease;
CC cardiovascular disease; diabetes mellitus;
CC erythematosus; severe combined immunodeficiency (SCID); AIDS; viral,
CC bacterial or fungal infection; malaria; autoimmune disorders; asthma,
CC allergies; aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antinephritic disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.

CC Sequence 6816 BP; 1977 A; 1523 C; 1557 G; 1757 T; 2 other;

alignment_scores:

Quality: 2244.50 Length: 2149
Ratio: 1.861 Caps: 56
Percent Similarity: 56.119 Percent Identity: 28.478

alignment_block:

US-09-737-246-2 x AAC74524 ..

Align seg 1/1 to: AAC74524 from: 1 to: 6816

```
94 ArgAspCysArgThrLeuValSerAlaValProGluGluSerGluMetAs 110
   ||||| : : : : ||| : : : : |||||
101 CAGAGCAAAAGTAGTATCTCTTGCGACGACAGACGTGAAGTGGAAATG... 148
110 ProHisValArgAspCysIleLeuSerTyrThrGluAspTyrPalaIleV 127
   ||||| : : : : ||| : : : : |||||
149 .....GAAAGATGATGCACAA 164
127 allleargLysTyrHisLysLeuGlyThrGlyPheAsnProAsnThrLeu 143
   : : : : : ||| : : : : : ||| : : : : :
165 TTCATAATAGATCTCCACCTCAAC..... 190
144 AspLysGlnLysGluArgGlnLysGlyLeuProLysGlnValPheGluSe 160
   ||||| : : : : ||| : : : : |||||
191 .....TTTGAAGC 198
160 r.....AspGluAlaProAspGlyAsnSerTyrGlnAspAspGlnAspA 175
   : : : : : ||| : : : : : ||| : : : : :
199 TGCATTCAGAAAGGAAATGCGACCTCTCAGAAAGATGATGAACAA. 247
175 spleuLysArgArgSerMetSerIleAspAspThrProArgGlySerTyr 191
247 ..... 247
192 AlaCysSerIlePheAspLeuLysAsnSerLeuProAspAlaLeuLeuP 208
   ||| : : : : : ||| : : : : : |||||
248 .....AGCAAAATTTGGAAAGTTCTGTGTCGGTTTAGATAGCTACCTGCC 291
208 oAsnLeuLeuAspArgThrProAsnGlnGluIleAspArgGlnAspAspA 225
   ||||| : : : : ||| : : : : |||||
292 GGAACCTGCCAAGAGTGCAGAGACAGACAAATC..... 325
225 spGlnArgLysSerAsnArgHisLysGluLeuPheAlaLeuHisProSer 241
   : : : : : ||||| : : : : : |||||
326 ..AAACTGAAAGTAGAAACAGAGATCAAACTTTTATTG.....GAC 367
242 ProAspGluGluProIleGluArgLeuSerValProAspIlePro... 257
   ||||| : : : : : ||| : : : : : |||||
368 CCAGATGCCAGAGCTTGACTTCTCATCGCTGAGCCAGACAGAACTGAAGTC 417
258 ....LysGluHisPheGlyGlnArgLeuValLysCysLeuSerLeu 273
   : : : : : ||||| : : : : : |||||
418 ATTGAAGAGAGAGTTGGAAAAAGATCTTGTCAAGTGCATGATATTAT 467
273 yspheGluIleGlu..... 277
   : : : : : ||| : : : : : |||||
468 CTTTCATTTGCAATGCTGTGTGCCGAAATGAAGAAAGACCCACTACA 517
278 ...IleGluProIlePheAlaSerLeuAlaLeuTyrAspValLysGly 293
   : : : : : ||||| : : : : : |||||
518 AATGTGAACCTTTCTTTGTACTCTATCCCTGTGTGACATAAAATCAAA 567
293 sLysLysIleSerGluAsnPheTyrPheAspLeuAsnSerGluGluMet 310
   : : : : : ||||| : : : : : |||||
568 CCGAAGATTTTCTCCATTTCCACGTAGACCTGAACCATTTCTCACTGA 617
310 ysgLysLeuLeuArgProHisValProProAla..... 320
   : : : : : ||| : : : : : |||||
618 GCGAAATGAT.CGCCACACACGTCCCGCGCTGATGAATGACAGTGGGCC 666
321 .....AlaIleThrThrLe 325
   |||
667 GAAACCAATCTGCCCTCAGGGGACATCTTCATGAAGCCGCCCATGCAGTA 716
325 uAlaArgSerAlaIlePheSerIleThrTyrProSerGlnAspValPhe 342
   : : : : : ||||| : : : : : |||||
717 TCCGAGAGGAGATATTTTACGACTTGTGCTCATCCAGATATATTC 766
342 euValIleLysLeuGluLysValLeuGlnGlnLysPheIleGlyLys 358
   ||||| : : : : : ||||| : : : : : |||||
```

```
767 TTGTGCCAGAAATTGAAAAAGTCTT...CAGGGAGACATCAACATTTGC 813
359 AlaGluProTyrMetIlePheLysGluAlaAspAlaThrLysAsnLysG 375
   ||||| : : : : : ||| : : : : : |||||
814 GCTGAGCCATATATG.....AAAAGTTCAGACTCTCTTAAGGTGGCCCA 857
375 uLysLeuGlnLysLeuLysSerGlnAlaAspGlnPheCysGlnArgLeu 392
   : : : : : ||| : : : : : |||||
858 GAAGTGTCTGAAG.....ATGCCAGACGACGATCCCAAGACGTAG 898
392 LysTyrTArgMetProPheAlaThrPheAlaIleHisLeuMetAsnIle 408
   ||||| : : : : : ||||| : : : : : |||||
899 GACACTATAGAAATGCCATTTGCTGGCAGACAGACATGTTGTAAG... 945
409 ValSerSerAlaGlySerLeuGlnArgAspSerThrGluValGluIleSe 425
   : : : : : ||||| : : : : : |||||
946 ..GATGCATCTGGAAATCTTGACAAAATGCC..... 975
425 rThrGlyLysArgLysGlySerTyrSerGluArgArgAsnSerSerIleV 442
   ||| : : : : : ||| : : : : : |||||
976 .....AGATTTCTGCGCATC. 990
442 aGlyArgArgSerLeuGluArgThrThrSerGlyAspAspAlaCysAsn 458
990 ..... 990
459 LeuThrSerPheArgProAlaThrLeuThrValThrAsnPhePheLysG 475
991 .....TACAGCA 998
475 nGluLysAspArgLeuSerAspGluAspLeuTyrLysPheLeuAlaAsp 492
   ||||| : : : : : ||||| : : : : : |||||
999 AGACAGCATATAGCTATCATCATGATGACATGCTCAAGTTACTGTGAGACT 1048
492 eTArgArgProSerSerValLeuArgArgLeuArgProIleThrAlaGln 508
   : : : : : ||||| : : : : : |||||
1049 TTCGGAAACCTGAGAG...ATGGCTAAGCTCCCGATATTTTAGGCAT 1095
509 LeuLysIleAspIle...SerProAlaProGluAsnProHisTyrCysLe 524
   ||||| : : : : : ||| : : : : : |||||
1096 CTAGACATTACAAATTGATATGTTTCTCAGACTTCCCTAATTATGTTAA 1145
524 uThrProGluLeuLeuGlnValLysLeuTyrProAspSerArgValArg 541
   : : : : : ||||| : : : : : |||||
1146 TTCATCATCATCTCCACACAAAACATTTGAAACCGTAGTAACCTCCCA 1195
541 rofThrArgGluIleLeuGluPhe...ProAlaArgAspValTyrVal... 555
   ||| : : : : : ||||| : : : : : |||||
1196 TCACGTTTGAAGTGGAGAAATTTGTGCCCTGCATCCAAACACACTCAG 1245
556 ProAsnThrThrTyrArgAsnLeuLeuTyrIleTyrProGlnSerLeuAs 572
   ||||| : : : : : ||||| : : : : : |||||
1246 CCTTACACCATCTACACCATCATCCTTACGTTATCCTAAGTACTTAA 1295
572 nPheAlaAsnArgGln.....GlySerAlaArgAsnIleThrValLys 587
   : : : : : ||||| : : : : : |||||
1296 ATACGACATCATCAAGAGTCTTTGCCAAGGCTGAAATATTTGGATTGCA 1345
587 aGlnPheMet...TyrGlyGluAspProSerAsnAlaMetProValIle 602
   : : : : : ||||| : : : : : |||||
1346 TTGAATTCAAAGATTCAGATGAGAAAGCTTCACCCCTTAAGTGCAT 1395
603 pheGlyLysSerSerCysSerGluPheSerLysGluAlaTyrThrAlaVa 619
   : : : : : ||||| : : : : : |||||
1396 TATGGCAGACTGTGTGGCCACTTTTCACAGAGAGCGCTTGTGCTGACGT 1445
619 lValTyrHisAsnArgSerProAspPheHisGluLysValLys 636
   : : : : : ||||| : : : : : |||||
1446 TTACACCATCAACAAACCAAGAAATTTATGATGAGATTAATAATAGAGT 1495
636 euProAlaThrLeuThrAspHisHisHisLeuLeuPheThrPheTyrHis 652
   ||||| : : : : : ||||| : : : : : |||||
1496 TGCCACTGCTGCTGATGAAGAAAGACACACTGTGTGCTACATTTTCAT 1545
```

```

653 ValSerCysGlnGln.....LysGlnAsnThrProLe 663
    |||||.....
1546 GTCAGCGTGTACAACTCAAGTAAGACAGACAGACAGAGGATCTCCT 1595
    |||||.....
663 uGluThrProValGlyTyrThrThrPleProMetLeuGlnAsnGlyArgL 680
    |||||.....
1596 TGAACCCAGTGTGGCTACTCTGGCTCCCTCCCTGAAAGACGGAAGG 1645
    |||||.....
680 euLysThrGlyGlnPheCysLeuProValSerLeuLulysProPro... 695
    |||||.....
1646 TGGTGACACGAGACACACATCCGCTCTCGGCAACCTTCCTTCGGGC 1695
    |||||.....
696 ...GlnAlaTyrSerValLeu.....SerProGluVal 705
    |||||.....
1696 TATCTTGGCTACCAAGAGCTTGGATGGCAGGACATTATGTCGGAAT 1745
    |||||.....
705 LProLeuProGlyMetLysTrpValAspAsnHisLysGlyValPheAsnV 722
    |||||.....
1746 T.....AAATGGGTAGATGAGGCAAGCCACTGCTGAAAA 1780
    |||||.....
722 aLgLuValAlaValSerSerLeHisThrGlnAspProGlyLeuAsp 738
    |||||.....
1781 TTTCACACTACTGTGTTCTACAGTGTATCTAGGATCAGCATTTTACAT 1830
    |||||.....
739 LysPhePheAlaLeuValAsnAlaLeuAspGluHisLeuPheProValAr 755
    |||||.....
1831 AATTTTTCAGTACTGTCAAGAAACCGAATCTGGA..... 1866
    |||||.....
755 gLlGgLyAspMetArgLLeMetGlnAsnAsnLeuGlnAsnGluLeuLys 772
    |||||.....
1867 .....GCCCAAGCCTTAGGAAACGAACTTGTA 1894
    |||||.....
772 eSerLeSerAlaLeuAsnSerSerGlnLeuGluProValValArgPhe 788
    |||||.....
1895 AGTACCTTAGAGCTGTGATGCGAGTGAAGCCACGATGATCGCTTC 1944
    |||||.....
789 LeuHisLeuLeuLeuAspLysLeuLeuLeuValLleArgProProVa 805
    |||||.....
1945 TTGCCCATATCCTTAACACAGCTGTCCGAGTCTCACCAGA..... 1986
    |||||.....
805 LlleAlaGlyGlnLleValAsnLeuGlyGlnAlaSerPheGluAlaMet 822
    |||||.....
1987 .....GCCACACAGAGAAAGATCG 2005
    |||||.....
822 LaseLleIleAsnArgLeuHisLysAsnLeuGluGlyAsnHisAspGln 838
    |||||.....
2006 CCGTTTACGTGACTCGGTCATTTATTCATGTGTTGCCAGTGCATGAG 2055
    |||||.....
839 HisGlyArgAsnSerLeuLeuAlaSerTyrLleHisTyrValPheArgLe 855
    |||||.....
2056 GAAGGATGTGAGAGCCATTGAGGTCAATGTTAGTACGCGTAT.... 2100
    |||||.....
855 uProAsnThrTyrProAsnSerSerSerProGlyGlyLeuGlyG 872
    |||||.....
2100 ..... 2100
    |||||.....
872 LysValHisTyrAlaThrMetAlaArgSerAlaValArgProAlaSer 888
    |||||.....
2100 ..... 2100
    |||||.....
889 LeuAsnLeuAsnArgSerArgSerLeuSerAsnSerAsnProAspLese 905
    |||||.....
2100 ..... 2100
    |||||.....
905 rGlyThrProThrSerProAspAspGluValArgSerLleGlySerL 922
    |||||.....
2100 ..... 2100
    |||||.....
922 ySGLyLeuAspArgSerAsnSerTrpValAsnThrGlyGlyProLysAla 938
    |||||.....
2101 .....AAGCT 2106
    |||||.....

```

```

939 AlaProTrpGlySerAsnProSerProSerAla.....GluSerThrGln 953
    |||||.....
2107 GAGCCATATGTTGGCTTGATACAGACACTGCATGACAGAACTGACCAA 2156
    |||||.....
953 nAlaMetAspArgSerCysAsnArgMetSerSerHisThrGluThrSers 970
    |||||.....
2157 ATCCATG.....ACCA 2167
    |||||.....
970 eRrPheLeuGlnThrLeuThrGlyArgLeuProThrLysLeuPheHis 986
    |||||.....
2168 CGATTCTCAAGCCTTTCGCGATTTCCTCACCCACCAACCACTA..... 2211
    |||||.....
987 GluGluLeuAlaLeuGlnTrpValValCysSerGlySerValArgGluSe 1003
    |||||.....
2211 ..... 2211
    |||||.....
1003 rAlaLeuGlnGlnAlaTrpPhePhePheGluLeuMetValLysSerMetV 1020
    |||||.....
2212 ...CTGAAGTACATCATGTTTCTTTGATGTACTGATCAATATCATG 2257
    |||||.....
1020 aLHisHisLeuTyrPheAsnAspLysLeuGluAlaProArgLysSerArg 1036
    |||||.....
2258 CTCACCATTTGATAGAGACATCAAAAGTTAAGTTGCTGCGAAACAGAGA 2307
    |||||.....
1037 PheProGluArgPheMetAspAspLleAlaAlaLeuValSerThrIleAl 1053
    |||||.....
2308 TTTCCTGATGCTTATCATCATGACAGTGAACCCCTTTAAATATGCTGAT 2357
    |||||.....
1053 aSerAspLleValSerArgPheGlnLysAspTrpGluMetValGluArgL 1070
    |||||.....
2358 GCCACACATCAGTCAAGAGTTTCAGATTAATTCAGAGCATCTGAAGACG 2407
    |||||.....
1070 euAsnThrSerLeuAlaPhePheLeuAsnAspLeuLeuSerValMetAsp 1086
    |||||.....
2408 CGAATCATATGCTTGTCTGTCTTCATCAAGAGATGTTTACCCTTCATGAC 2457
    |||||.....
1087 ArgGlyPheValPheSerLeuLleLysSerCysTyrLysGlnValSerSe 1103
    |||||.....
2458 AGGGCTTTGTCTTC.....AAGCAGATCAACAA 2486
    |||||.....
1103 rLysLeu.....TyrSerLeuProAsnProSerValLeuValSerLeuAl 1118
    |||||.....
2487 CTACATTTAGCTGTTTGTCTGCTGAGACCCAAAGACCTCTTTGAATACA 2536
    |||||.....
1118 rGlyLeuAspPheLeuArgLleLleCysSerHisGluHisTyrValThrLeu 1134
    |||||.....
2537 AGTTTGATTTTCTCGGTAGTGTGCAACCATGAACATTAATTTCCCTG 2586
    |||||.....
1135 AsnLeuPro...CysSerLeuLeuThrProProAlaSerProSerProSeV 1151
    |||||.....
2587 AACTTACCAATGCCATTTG..... 2605
    |||||.....
1151 aLSerSerAlaThrSerGlnSerSerGlyPheSerThrAsnValGlnAsp 1167
    |||||.....
2606 .....GAAAGGCAAGATTC.....AAGAT 2626
    |||||.....
1168 GlnLysLleAlaAsnMet...PheGluLeuSerValProPheArgGlnGln 1183
    |||||.....
2627 ACCAAGACCTCCAGCTTGACTACTCATTAACAGATGAGTTCTCGAGAAC 2676
    |||||.....
1184 HisTyrLeuAlaGlyLeuValLeuThrGluLeuAlaValLleLeuAspPr 1200
    |||||.....
2677 CACTTCTTGGGAGCTGTACTAGGAGAGTGGGACACCCCTCCACAGA 2726
    |||||.....
1200 oAspAlaGluGlyLeuPheGlyLeuHisLysLysValLleAsnMetValH 1217
    |||||.....
2727 GTTCCGGGAG.....GTCCGTGATCGCCATCAGTGTGCTCA 2764
    |||||.....
1217 iAsnLeuLeuSerSerHisAspSerAspProArgTyrSerAspProGln 1233
    |||||.....
2765 AGAACCTGTGATTAAGCATTTTGTGATGACAGATATGCTCAAGAGC 2814
    |||||.....
1234 lLleLysAlaArgValAlaMetLeuTyrLeuProLeuLleGlyLleIleMe 1250
    |||||.....

```

```

2815 CATGAGGCAAGATAGCCACCTCTACCTGCTGTTGGCTGCTGAT 2864
1250 tGluThrValProGlnLeu.....TyrAspPheThrGluThrHisAsnG 1265
2865 TGAAAACGTCACGGGATCAATGTGAGGAGTGTGCACCTTCCTCCGTGCA 2914
1265 InArgGlyArgProIle..... 1270
2915 ACCGGGCGATAGCTGTGAAGATGAATCCCTGCGCTACACAGCTGTGAAT 2964
1271 CysIleAlaThrAspAspTyrGluSerGluSerGlySerMetIleSerG1 1287
2965 CCGCTGTGAGCGCCGACAGAGGAGACCTGACACACACCTGCACAA 3014
1287 nThrValAlaMetAlaIleAlaGlyThrSerValProGlnLeuThrArgP 1304
3015 GGACCTGTGGGGCCATCTCGGCATTGCTTCATATACAAACCTCA 3064
1304 ro.....GlySerPheLeu 1308
3065 CTCCAAACATCAACAGTGTGAGAAATGCTGATCGAGAGGATCT...CTC 3111
1309 LeuThrSerThrSerGlyArgGlnHisThrThrPheSerAlaGluSerSe 1325
3112 ATAAACACAGATTGCGGTAAACAGCTTCACAAAGAAATAGTAGAAGAG 3161
1325 r..... 1325
3162 CAATTCCCTGATAAGACAAACAAAGTAGACATTTGGAAATCCGTGG 3211
1326 .....ArgSerLeuLeuIleCys 1331
3212 TTCGCTGTGATAACTGTGACCGCTGTGAGATTAAAGCCTACTGATGCT 3261
1332 LeuLeuThrPValLeuLeuAsnAlaAspGluThrValLeuGlnLysTrpPh 1348
3262 TTCCTCTACATCTTAAAGAGCATGCTGATGATGCTTTGTTACTATATG 3311
1348 eThrAspLeuSerValLeuGlnLeuAsnArgLeuLeuAspLeuLeuTyrL 1365
3312 GAACAGGCTTCACATCTGAACCTATGATTTTATTCACATATCTGAAG 3361
1365 euCysValSerCysPheGluTyrLysGlyLysValPheGluArgMet 1381
3362 TCTGCTGTGACCACTTCACATGAGGGAACGATACATACCCAGAACCA 3411
1382 AsnSerLeuThrPheLysLysSerLysAspMetArgAlaLysLeuGluG1 1398
3412 GGAAATG.....ATGCATGCCAGATTGCACCA 3437
1398 uAlaIleLeuGlySerIleGlyAlaArgGlnLumetValArgArgSera 1415
3438 G.....CTGGGCACCTG..... 3450
1415 rgGlyGlnLeuGlnArgSerProSerGlySeraLAPheGlySerGlnGlu 1431
3451 .....GATTAAC 3456
1432 AsnLeuArgTrpArgLysAspMetThrHisTrpArgGlnAsnThrGluL 1448
3457 TCTCTACATTTTAACACAGCTATGCGCAC..... 3486
1448 sLeuAspLysSerArgAlaGluIleGlnHisGlnAlaLeuIleAspGlyA 1465
3487 .....TCGGACGACAGATGTTCTGCACACAGCATTTACTTGAAGCA 3526
1465 snLeuAlaThrGlnAlaAsnLeuIleIleLeuAspThrLeuGlnLueVal 1481
3527 ACATTCCTCTAGAGGTTCCCTGACAGCTGTGACACACGCTTTCTATATT 3576
1482 Val.....GlnThrValSerValThrGluSerLysGluSerI1 1494
::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
3577 ACATTGGCGTTTAAAGAACAGCTCTGGCCGACCATGACATTAATCCTCT 3626
1494 eleuGlyGlyValLeuLysValLeuLeuHisSerMetAlaCysAsnGlns 1511
3627 CATGAAAAAGATTTTGATGTCTACTCTGTGTTTCTTTCAAAAACATCACT 3676
1511 eAlaValTyrLeuGlnHisCysPheAlaThrGlnArgAlaLeuValSer 1527
3677 CTGAACAGGCTTTAAAAATGTCTCACTCTGCTTAAGAGTCCTTAATTATT 3726
1528 LysPheProGluLeuLeuPheGluGluGluThrGluGlnCysAlaAspLe 1544
3727 AAGTTTCCTCAACATTCTATGAAGGAGAGCCGACATGTGTGCGGCTCT 3776
1544 uCysLeuArgLeuLeuArgHisCysSerSerSerIleGlyThrIleArgS 1561
3777 GTGTTACGAAATTTCTAAGTGTCTAATCTCAAGCTGAGCTCATAGGA 3826
1561 eHisProSerAlaSerLeuLeuLeuMetArgGlnAsnPheGluIle 1577
3827 CGGAGGCTCCCACTCTCTACTCTGATGAGCAACAACTTGATTAC 3876
1578 .....GlyAsnAspPheAlaArgValLysMetGlnValProMetSerIe 1592
3877 ACTGAAAGAAGTCCTTGTCCGACACATTGTCAAGTCAATCATCTGT 3926
1592 uSerSerLeuValGlyThrSerGlnAsnPheAsnGluGluPheLeuArgA 1609
3927 CAGCCAGCTATAGCAGACGTTGTGGCATTTGGGAGACCAATTCACAG 3976
1609 rgSerLeuLysThrIleLeuThrTyrAlaGluGlnLysLeuGluLeuArg 1625
3977 AGTCCCTGTCAATCAACAACTGTGCCAACAGTAGCCGCTATTATAG 4026
1626 GluThrThrPheProAspGlnValGlnAspLeuValPheAsnLeuHis 1642
4027 CACACGAGCTTCTCCTGTGATGAAAGAGACTTAAACAAAGATAGCAG 4076
1642 tIleLeuSerAspThrValLysMetLysGlnHisGlnLysProGlnLum 1659
4077 GGTCTATATGGCCACCGCCAGATGAAGAGAGCATGGAAGACCCAGAGA 4126
1659 eLeuIleAspLeuMetTyrArgIleAlaLysGlyTyrGlnThrSerPro 1675
4127 TGTGTGTGACCTCCATACAGCTGTGCCAAATTCATGTCCACAGAGCCC 4176
1676 Glu...ArgLeuThrTyrLeuGlnAsnMetAlaGlyLysHisSerGluAr 1691
4177 GAGCTCAGGAAGAGAGTGCTGCACAGCATGTGCCAGATGTCATAAAA 4226
1691 gSerAsnHisAlaGlnAlaAlaGlnCysLeuValHisSerAlaAlaLeuV 1708
4227 TGGGATCTCTCAGAGGACGCAATGTCTATGTCCAGTAAACAGCCCTAG 4276
1708 aAlaGluTyrLeuSerMetLeuGlnAspArgLysTyrLeuPro..... 1722
4277 TGGCAGAATATCTCACACGGAAGAAACACAGTCCAGTGGAGCCGCCCTT 4326
1722 ..... 1722
4327 CTCGCCACAGCATAGCGCTGCTGAGGAGAGCGGGAGGCGTGT 4376
1723 ...ValGlyCysValThrPheGlnAsnIleSerSerAspValLeuGluG 1738
4377 TAGACAAGAGATGCAACGCTTCAGGGTCAATTAACCAACATGAGAGAG 4426
1738 LuSerAlaValSerAspAspValValSerProAspGluGluGlyIleCys 1754
4427 AGCGCTCCATGATGAGAGAGAGCTGGGATGAGAGAT..... 4461
1755 SerGlyLysTyrPheThrGlnArgSerGlyLeuValGlyLeuGlnGlnAl 1771
4462 .....GTCCATTTCAACGAGAGATGTGCTGATGAGACTCTCTGAGCAGTG 4505

```



```

293 stlyslsileserGIuAsnPhetYrPheAspleu.....Asnserg 307
   ::::::::::::::::::::|||
365 GTTGGAGAGTGGCGCACATGATCTATGACCTTATTGAATGGCGATCAC 414
   ::::::::::::::::::::|||
307 IuGIuMetLysGIuLeuLeuArgProHisValProProAlaAlaIleThr 323
   ::::::::::::::::::::|||
415 AAATTCTTCTGGAACCTCG.....CCTCAGATGAACTCAAA 452
   ::::::::::::::::::::|||
324 ThrIleuAlaArgSerAlaIlePheSerIleThrTyProSerGIuAspVa 340
   ::::::::::::::::::::|||
453 GAAGTGAAGAAAGAGGTCAACCCAAAATGATTATGGAACAGA..... 497
   ::::::::::::::::::::|||
340 IPheuValIleLysLeuGIuLysValLeuGIuInGIuAspIleGIu 357
   ::::::::::::::::::::|||
498 .....ATTCTAGATTGGACCTGGTGTGATGAGATGAGAT...GGGA 537
   ::::::::::::::::::::|||
357 IuCyAlaGIuProTyMet.....IlePheLys..... 366
   ::::::::::::::::::::|||
538 ATATTGTTGGATCCAGAAATTAACTAGACAGATTAGTCTCTCAGAGCTCAT 587
   ::::::::::::::::::::|||
367 GIuAlaAspAlaThrLysAsnLysGIuLysLeuGIuLysLysSerGI 383
   ::::::::::::::::::::|||
588 GAAATAGCTTTTAAACAAGTGGAGGAAAGGTTTACAAAGAGAAATCTCA 637
   ::::::::::::::::::::|||
383 nAlaAspGlnPhe...CysGlnArgLeuGIuLysTy...ArgMetProp 398
   ::::::::::::::::::::|||
638 AAAGCAGAAACATAGATATTAAACAGCAAGCCAGTTGCTGCAACCCCTT 687
   ::::::::::::::::::::|||
398 hAlaITrPThrAlaIleHisLeuMetAsnIleValSerSerAlaGIuSer 414
   ::::::::::::::::::::|||
688 CTCTGGCCCTTGTGTGAACCTCAAAAATGTGTTGTAAATAGGA... 734
   ::::::::::::::::::::|||
415 LeuGIuArgAspSerThrGIuValGIuLysSerThrGIuLysGIuArgL 431
   ::::::::::::::::::::|||
735 .....GAAGATGCTGAAGTCTCATGTCTCATATGACCCCTGCGA 775
   ::::::::::::::::::::|||
431 ySer.....TrpSerGIu..... 435
   ::::::::::::::::::::|||
776 GTCCAAATTTCATCATGAGAACTACCTGTTGCGTGCATGTCAGAT 825
   ::::::::::::::::::::|||
436 .....ArgArgAsnSerSerIle 441
   ::::::::::::::::::::|||
826 TACCTAAAGACATAGACAGATTACATTAATTGCCAGCCGTGTTACTGAC 875
   ::::::::::::::::::::|||
442 ValGIuArgSerLeuGIuArg..... 449
   ::::::::::::::::::::|||
876 CTCGAGACAAAGACCTGAAAAGGAGAAATCAGTTGTTCTGTGCAGAT 925
   ::::::::::::::::::::|||
450 .....ThrThrSerGIuAspAspAlaCysAsnLeu 460
   ::::::::::::::::::::|||
926 TGTTCGCGTGGGTGCATGAGCTGAGGACACAAACACAGAACTGA 975
   ::::::::::::::::::::|||
460 hSer.....PheArgProAlaThrLeuThrValIleThrAsn 471
   ::::::::::::::::::::|||
976 CTTCCGGGTGGCGGACCTTTTGGAGTGGCTGTGATGATGAACAGAT 1025
   ::::::::::::::::::::|||
472 PhePheLysGIuInGIuLysArgIleuSerAspGIuAspLeuTyLysPh 488
   ::::::::::::::::::::|||
1026 ATAAATA.....AATGAAAAGTAGATGATGAAGATAAAGACAGATT 1066
   ::::::::::::::::::::|||
488 eLeuAlaAspMetArgArgProSerSerValLeuArgArgLeuArgPro 505
   ::::::::::::::::::::|||
1067 CATTCGCC.....TTTCAGCCGG 1083
   ::::::::::::::::::::|||
505 IeThrAla.....GlnLeuLysIleAspIleSerProAla 516
   ::::::::::::::::::::|||
1084 TGGCAGGGGAGATGACTTCTCTCAGACTGTATTAACAAGTCAATCCCT 1133
   ::::::::::::::::::::|||
517 ProGIuAsnProHisTyCysLeuThrProGIuLeuLeuGIuValLysLe 533
   ::::::::::::::::::::|||
1134 GCCAAAGAGTCAACCCAAAGGGGAGGTTGTGGTAAACATTGGAATT 1183
   ::::::::::::::::::::|||
533 uTyArgProAspSerArgValArgProThrArgGIuIleLeuGIuPhePro 549

```

```

1184 ACTTCCCTGGAGAT...ATCCATCAGATCCGAAAA.....GAOTTCGCC 1224
   ::::::::::::::::::::|||
550 .....AlaArg.....Asp 552
   ::::::::::::::::::::|||
1225 ATTAGTGGACAGACACACAGCTGTGGCTGAAAAAAGAGGTTTCCGGAG 1274
   ::::::::::::::::::::|||
553 ValTyValProAsnThrThrTyArgAsnLeuLeuTyIleTyProGI 569
   ::::::::::::::::::::|||
1275 ATATATCTGCTGCTGATGT...CGAAATGATATCTATGTAACATTAGT 1321
   ::::::::::::::::::::|||
569 nSerLeuAsnPhe...AlaAsnArgInGIuSerAlaArgAsnIleTy 585
   ::::::::::::::::::::|||
1322 TCAAGAGATTTTGTATTAAGAAAGCAAAACACAGGAGAAAGCTGGAG 1371
   ::::::::::::::::::::|||
585 aLysValGlnPheMetTyGIuAspProSerAsnAlaMetProVal 601
   ::::::::::::::::::::|||
1372 TCACGGTGTCT...GTCATAGATAGAGGATGGAAACGATTAGACATGTG 1418
   ::::::::::::::::::::|||
602 IlePhe.....GIuLysSerSerCysSerGIuPheSerLysGIuAl 615
   ::::::::::::::::::::|||
1419 ATTTTCCGGGTCTGTGTATGAAAGCATTTTCAGAG..... 1454
   ::::::::::::::::::::|||
615 aTyThrAlaValValTyHisAsn...ArgSerProAspPheHisGIu 631
   ::::::::::::::::::::|||
1455 .TACAAATCTGTGATTACTTACCAAGTAAAGCAAGCCAGCTGTTGAGA 1503
   ::::::::::::::::::::|||
631 IuIleLysValLysLeuProAlaThrLeuThrAspHisHisLeuLeu 647
   ::::::::::::::::::::|||
1504 CTGTAAAGTGGCCATTCCCATCGAGAGCGTAAACCGCATCTTCGG 1553
   ::::::::::::::::::::|||
648 PheThrPheTyHisValSerCysGln...GlnLysGlnAsnThrProle 663
   ::::::::::::::::::::|||
1554 TTTACCTTCGCCCAAGGTATCACAAGACTCTAAAGATTAATCTGAGAA 1603
   ::::::::::::::::::::|||
663 uGIuThrProValGIuTyThrThrPheProMetLeuGlnAsnGIuArgL 680
   ::::::::::::::::::::|||
1604 AATATTGCACTGATGATTTGTCAAGCTGATGATGATGATGATGATGAT 1653
   ::::::::::::::::::::|||
680 eulstThrGIuGlnPheCysLeuProVal.....SerLeuGIuLys 693
   ::::::::::::::::::::|||
1654 TGGCAGCGGAGAGACAGATCTTATGCTATTAAGCCGAGGAGGAAG 1703
   ::::::::::::::::::::|||
694 ProProGlnAlaTySerValLeuSerProGIuValProLeuProGIu 710
   ::::::::::::::::::::|||
1704 CTGAGAGATCTGCCACAGTACTTGAGT.....CTGCCCTCCAC 1741
   ::::::::::::::::::::|||
710 tLysTrPValAspAsnHisLysGIu..... 718
   ::::::::::::::::::::|||
1742 GAAGGAGAGATTGGAAGAAAGGCCACCTCGGCCAGCGCAAGAGCATCG 1791
   ::::::::::::::::::::|||
719 .....ValPheAsnValGIuVal 724
   ::::::::::::::::::::|||
1792 AGAGCCCTTGGAGCTCACCATTAGCAAGAGACTCTTCCAGATCTCCAG 1841
   ::::::::::::::::::::|||
725 ValAlaValSerSerIleHisThrGlnAspProTyLeuAspLysPheh 741
   ::::::::::::::::::::|||
1842 CTGCTGTCTCCACCAAACTGACTCAGAACGTGGACCTT..... 1880
   ::::::::::::::::::::|||
741 eAlaLeuValAsnAlaLeuAspLuhHisLeuPheProValArgIleGI 758
   ::::::::::::::::::::|||
1881 .....CTGGGCC 1887
   ::::::::::::::::::::|||
758 spMetArgIleMetGIuAsnAsnLeuGIuAsnGIuLeuLysSerSerIle 774
   ::::::::::::::::::::|||
1888 TCTTGAATGGCGCTCCAAACAC.....AGCTGCTCAGAGAGACTGG 1931
   ::::::::::::::::::::|||
775 SerAlaLeuAsnSerSerGIuLeuGIuProValValArgPheLeuHis 791
   ::::::::::::::::::::|||
1932 AGGCAGCTGATGAAGTCAATGCTGTGTAAGTGAATTTCTTCAAGA 1981
   ::::::::::::::::::::|||
791 uLeuLeuAspLysLeuIleLeuLeuValIleArgProProValIleAla 808
   ::::::::::::::::::::|||

```


3663	..GAAATGATATAAGGATTTGTACAG..	3669
1674	SerProgluArgLeuThrTrpLeuGlnAsnMetIaGlyLysHisSerGlu	1690	
3690	CTCTGTGACCTGCACACAGA	3709
1690	uArgSerAsnHisIaGluIaIaGlnCysLeuValHisSerIaIaIaL	1707	
3710	GTGTGATACTACACGACGGGCTTACACCTTGCTT.....	C	3747
1707	euValIaIaGluTrpLeuSerMetLeuGlnAspArgLysTyrLeuProVal	1723	
3748	TTCATCAAGGCTTCTTAAGTGTCGAGAGAT.....		3779
1724	GlyCysValThr.....	PheGlnAsnIleSerSer	1733
3780	GTGTGTGTGGCCACCTACCCACGGCGAGGGTACCCAGCCACACGCA	3829	
1733	AsnValIleuGlnGluSerIaIaValSerAspAspValIaSerProAsp.	1749	
3830	GGGACAGCTGAAGAGACACCTCTACACAGAAATCATCATTCATTCGACA	3879	
1750GluGlnGlyIleCysSerGlyLysTyrPheThrGlu	1761	
3880	AAGCAAGATGTGGGAGGAGGCATTCGCTTGGCGCAG.....	3917	
1762	SerGlyLeuValGlyLeuLeuGlnIaIaIaIaIaSerPheSerMetAl	1778	
3918	GAGCTAC	3925
1778	aglyMetTyrGlnIaIaValaAsnGluValTyrLysValLeuIleProIleH	1795	
3926	CGAGCAGTATGAG.....AAGCAATGTTT.....	3950	
1795	IsoGluIaAsnArgAspAlaLysLysLeuSerThrIleHisGlyLysLeu	1811	
3951GATTATGAGCAACTCAGGAAATTCCTGAAAAGAG	3966	
1812	GlnGluIaIaPheSerLysIleValHisGlnSerThrGlyTrpGluArgMe	1828	
3987	GCTCAGCTTTATGAACACATCTGTCAAAGGATC.....	AGGCC	4024
1828	TrpGlyThrTyrPheArgValaGlyPheTyrGlyThrLysPheGlyAsp.	1844	
4025	CAGCCTGACTATTTGCTGTGGCTACGACAGCAAGGGTCCCCACAT	4074	
1845	..LeuAspGluGlnGluPheValTyrLysGluProAlaIleThrLysLeu	1860	
4075	TCTCTGGGGGAAAAGTTTTCATTTCACGGGAAAGATAGAGCCCCG	4124	
1861	AlaGluIleSerHisArgLeuGlnGlyPheTyrGlyGluArgPheGlyGlu	1877	
4125	GAGGATTTTGAAGCTCGGCTTCTTAACCTACGTTCCAAAC.....	4163	
1877	uAspValIaGluValIleLysAspSerAsnProVal...AspLysCysL	1893	
4164GCCGAGAAATGAACACAAACATCTCCACAGGCGACGATATTA	4206	
1893	ysLeuAspProAsnLysAlaTyrIleGlnIleThrTyrValaGluProTyr	1909	
4207	AAACCTCTCTGCGCAG..TATATTCAGTGGCTTCACAGTGAAGCCCAA	4253	
1910	PheAspThr.....TyrGluMetLysAspArgIleThrTyr	1921	
4254	CTGCAATCTGCTCTTAAGTTTACAGGCGCAGTGTGAAGCAGATTTGAG	4303	
1921	TrpAspArgLysAsnTyrAsnLeuArgArgPheMetTyrCysThrTrpProPhe	1938	
4304	TTTTTACAGGCGTACAGAGCTCCAGCGATTTGAATATTTCTGGGCCAAATCC	4353	
1938	hIleu...AspGlyArgAlaHisGlyGluLeuHisGluGlnPheLysArg	1953	

```

4354 GGAGGGAGAGAAAAACCCAGACATGATTTGGCAATATGTGATGAG 4403
1954 LysThrIleLeuThrThrSerHisAlaPheProTyrIleLysThrArgVal 1970
      ::::: ||||| ::::: |||||
4404 AGAACATATATACACATGATATAAATTACCTGGAATTTAAGTGCTT 4453
1970 LAsnValThrHisLysGluGluIleIleLeuThrProIleGluValAlaI 1987
      ::::: ||||| ::::: |||||
4454 TGAGGTCAAGTCTGTTTCATGATGGAAATACGCCCTCGAGATGCCA 4503
1987 LeGluAspMetGlnLysLysThrGlnGluLeuAlaPheAlaThrHisGln 2003
      ||||| ||||| ::::: |||||
4504 TCGAGACCATGACCTGACGACAGACAGATCAACAGCATGCTGACAG 4553
2004 AspProAlaAspProLysMet.....LeuGlnMetValIleuGlu 2016
      ||||| ::::: |||||
4554 CACCTGATGACCCACGCTGCCATCAACCCGCTCTCCATGCTCTTAA 4603
2016 nGlySerValGlyThrThrValAsnGlnGlyProLeuGluValAlaGlnV 2033
      ::::: ||||| ::::: |||||
4604 CGGCATGCTGACCCACGCTGATGAGGGGCTGCGCAAACTACGAAAG 4653
2033 alphaLeuSer.....GluIleProSerAspProLys 2036
      ::||| :::::
4654 CCTCTTACAGACCGTACTCGACGAGACCCCTGAGCCCATGAAAG 4703
2037 .....GluIleProSerAspProLys 2043
4704 ATCGAAGAGCTCAGAGACCTGCTGCGAGATTCCTTTTCGGCCGA 4753
2043 sLeuPheArg...HisHisAsnLysLeuArgLeuGlyPheLysAspPheT 2059
      ::||| ||| ::||| ::|||
4754 AGGATCAGAAATCATCGAGACAAAGTCACGAGGACTGAGGCCGTTCC 4803
2059 hLysArgCysGlnAspAlaLeuArgLysAsnLysSerLeuIleGlyPro 2075
4804 ACGAAGAGTAGGAGCGCTGTTTCAACAGCTGAAGAA.....AAG 4844
2076 ValGlnLysGluTyr 2080
      ||||| ::|||
4845 GTGAGAAAGACTAC 4859

```

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:us-08-971-988-1

seq_documentation_block:

```

: sequence 1, Application US/08971988
: Patent No. 5786461
: GENERAL INFORMATION:
: APPLICANT: MICHIOKI MATSUDA et al.
: TITLE OF INVENTION: CDNA OF DOCK180 GENE AND DOCK180 PROTEIN
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wenderoth, Lind & Ponack
: STREET: 805 Fifteenth Street, N.W., #700
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
: OPERATING SYSTEM: MS-DOS
: SOFTWARE: Wordperfect 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/971,988
: FILING DATE: 17-NOV-1997
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/588,985
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Warren M. Cheek, Jr.
: REGISTRATION NUMBER: 33,367

```

```

: REFERENCE/DOCKET NUMBER:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-371-8850
: TELEFAX:
:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 6519 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHEICAL:
: ANTI-SENSE:
: FRAGMENT TYPE:
: ORIGINAL SOURCE: spleen cell of homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 24..5619
: US-08-971-988-1

```

alignment_scores:

Quality	Ratio	Length	Gaps
242.50	0.287	2005	101

Percent Similarity: 42.195 Percent Identity: 19.352

alignment_block:

US-09-737-246-2 x US-08-971-988-1 ..

Align seq 1/1 to: US-08-971-988-1 from: 1 to: 6519

```

277 GluIleGluProIlePheAlaLeuAlaLeuTyrAspValLysGlu 293
      ||| ::||| ::|||
315 GAGTGTCACCATCTGAGGACGCTCTACGTCAAGATACAGGAGAT 364
293 sLysLysIleSerGlnAspPheTyrPheAspLeu.....AsnSerG 307
      ::||| ::|||
365 GTTTCGAAGTGTGGGACATGATCTATGACTTATGATGCGCATAC 414
307 LucIleMetLysGlyLeuLeuArgProHisValProProAlaAlaIleMr 323
      ::||| ::|||
415 AAATCTTCTGTGAACTCTG.....CCTCAGATGACACAA 452
324 ThrLeuAlaArgSerAlaAlaIlePheSerIleThrTyrProSerGlnAsp 340
      ||| ::|||
453 GAAGTGAAGAGAGAGTACACGCCAAATGATGAGAAACAGA..... 497
340 lPheLeuValIleLysLeuGlnLysValLeuGlnGlnGlyAspIleGlyG 357
      ::||| ::|||
498 .....ATTCTGATTTGGACCTGCTGTAGACATGAAGAT...GGGA 537
357 LucYsAlaGluProTyrMet.....IlePheLys..... 366
      ::||| ::|||
538 ATATTTGGATCCAGATTAACAGACGATTAAGTCTCTTCAAGCTCAT 587
367 GluAlaAspAlaThrLysAsnLysGlnLysLeuGlnLysLysSerG 383
      ||| ::|||
588 GAATTAAGTCTTAAACAGAGTGAAGAGGTTACAAAGAGAAAAATCTCA 637
383 nAlaAspGlnPhe...CysGlnArgLeuGlnLysTyr...ArgMetProP 398
      ::||| ::|||
638 AAAGCAGACATAGATATTAAACAGACAGCAAGTTGCTGCAACCCCTT 687
398 heaIatPThrAlaIleHisLeuMetAsnIleValSerSerAlaGlySer 414
      ::||| ::|||
688 CTCTGGCCTTGTGTGACACCTCAAAAATGTGTTGTAATAATAGAA... 734
415 LeuLunArgAspSerThrGluValGluIleSerThrGlyLunArgLysG 431
      ::||| ::|||
735 .....GAGATGCTGAAGTCTCATGTCTCTATATGACCCCTGTGA 775
431 ySer.....TipSerGlu..... 435

```

```

776 GTCCAAATTCATCAGTAGAGACTACCTGGTTCGTCCAGTTCAGGAT 825
436 ..... ArgArgAsnSerSerIle 441
826 TACCTTAAGACATAGACAGATTACATTAATTTCCGAGCCGTGTTTACTGAC 875
442 ValGlyArgArgSerLeuGluArg..... 449
876 CTCGGAGACAAAGACCTGAAAGGAGAAATCAGTTTGTCTGTCCAGAT 925
450 ..... ThrThrSerGlyAspAspAlaCysAsnLeuT 460
926 TGTTCGCGTGGGTCCGATCGACGTGAGGCAACAACACAGAAATCGA 975
460 hSer..... PheArgProAlaThrLeuThrValThrAsn 471
976 CCTCGGGGTTCGGCGGACCTTTTGGAGTGGCTGTGATGGATGTACAGAT 1025
472 PhePheGlyGlnGluGlyAspArgLeuSerAspGluAspLeuTyrLysPh 488
1026 ATAATA..... AATGAAAGTAGATGATGAAGATTAAGCAGACATT 1066
488 eLeuAlaAspMetArgArgProSerSerValLeuArgArgLeuArgProI 505
1067 CATCC..... TTCAGCCGG 1083
505 LeThrAla..... GlnLeuLysIleAspIleSerProAla 516
1084 TGGCAGGGAGAGATGACTCTTCAGACTGTATTAAACAAAGTCATCCCT 1133
517 ProGluAsnProHisTyrCysLeuThrProGluLeuGlnValLysLe 533
1134 GCCAAAGAGTCAACCCACCAAGGGGCTTGTGGTAACTTAATAATT 1183
533 uTyrProAspSerArgValArgProThrArgGlnIleLeuGlnPhePro. 549
1184 ACTTCCTGGAGAT... ATCCATCAGATCCGAAAA..... GAGTTTCCGC 1224
550 ..... AlaArg..... Asp 552
1225 ATTTAGTGACAGACACACAGCTGTGCTCGAAAAACAGGGTTTCCGGAG 1274
553 ValTyrValProAsnThrThrTyrArgAsnLeuLeuTyrIleTyrProG 569
1275 ATAAATCATCCTCGTGATGTT... CGAAATGATATCTATGTACATTAGT 1321
569 nSerLeuAsnPhe... AlaAsnArgGlnGlySerAlaArgAsnIleThrV 585
1322 TCAGAGAGATTTTGTATTAAGAAACAAACACAGCAAGAACGTGAGG 1371
585 aLysValGlnPheMetTyrGlyGluAspProSerAsnAlaMetProVal 601
1372 TCACGGTGCT... GTGTACGATGAGGATGGGAAACGATTAGAGCATGTG 1418
602 IlePhe..... GlySerSerCysSerGluPheSerLysGlnAl 615
1419 ATTTTCCCGGGTCTGCTGATGATGAACGATTTCAGAG..... 1454
615 aTyrThrAlaValValTyrHisAsn... ArgSerProAspPheHisGln 631
1455 TACAATCTGTGATTACTACCAAGTAAAGCAGCCAGCGTGGTTGAGA 1503
631 LuIleLysValLysLeuProAlaThrLeuThrAspHisHisIleLeu 647
1504 CTGTTAAGGTGGCATTCCTCCATCGAGAGAGGATTAAACCGACGACCTTGG 1553
648 PheThrPheTyrHisValSerCysGln... GlnLysGlnAsnThrProLe 663
1554 TTACCTTCGCGCACAGTCATCACAGACCTTAAGGATTAATCTGAGAA 1603
663 uGluThrProValGlyTyrThrTyrIleProMetLeuGlnAsnGlyArgL 680
.....
1604 AATATTGCACTAGCATTTGTCAAGCTGATAGATACATGATACCAACC 1653
680 eLysThrGlnGlnPheCysLeuProVal..... SerLeuGlnLys 693
1654 TGCAGACGGAGACACAGATCTTATGCTGTAAAGCCGAGGAGGAAAG 1703
694 ProProGlnAlaTyrSerValLeuSerProGluValProLeuProGlyme 710
1704 CTGGAAGATGCTGCCACGACTTACTGAT..... CTGCCCTCCAC 1741
710 tLysTrpValAspAsnHisLysGly..... 718
1742 GAAGCAGAGATTGGAGAAAAAGGCCACTCGCCACCGGCAAGAGCATGC 1791
719 ..... ValPheAsnValGluVal 724
1792 AGACCTTGGGAGCTGCACCATTAAGCAAGACGTCCTCCAGATCTCCAG 1841
725 ValAlaValSerSerIleHisThrGlnAspProTyrLeuAspLysPhePh 741
1842 CTCGTGTGCTCCACCAACAGTCACTCAGAACGTGACCTT..... 1880
741 eAlaLeuValAsnAlaLeuAspGluHisLeuPheProValArgIleGlyA 758
1881 ..... CTGGGGC 1887
758 sPheTArgIleMetGluAsnAsnLeuGluAsnGlnLeuLysSerSerIle 774
1888 TCTTGAATGGCGCTCCACACAC..... AGCCTGCTGCAGGAGAACTTG 1931
775 SerAlaLeuAsnSerSerGlnLeuGluProValValArgPheLeuHisLe 791
1932 AGGCAGCTGATGAAGTCGATGCTGGTGAAGTAGTAAGATTGTTTCAGCA 1981
791 uLeuLeuAspLysLeuIleLeuLeuValIleArgProProValIleAlaG 808
1982 CAGCTTGATGCCCTTTCAAACATCATGATGAG..... AACT 2019
808 LysGlnIleValAsnLeuGlyGlnAlaSerPheGluAlaMetAlaSerIle 824
2020 CAGAGAGTGAAGACTTTTGACACGTTACTGCTTTGATGCTGCTGATTATTC 2069
825 IleAsnArgLeuHisLysAsnLeuGluGlyAsnHisAspIleHisGlyArg 841
2070 ATTGCAGT..... ATTCGTGATGAAGAAATTTTCAGCAT... TT 2104
841 gasnSerLeuLeuAlaSerTyrIleHisTyrValPheArgLeuProAsnT 858
2105 TAATCCTGTTTGGAAACTTACATTAAGAAACACTTACG..... 2144
858 hTyrTrProAsnSerSerSerProGlyProGlyLeuGlyGlySerVal 874
2144 ..... 2144
875 HisTyrAlaThrMetAlaArgSerAlaValArgProAlaSerLeuAsnLe 891
2145 ..... GCAACGTTAGCTTACAGCAACTTG..... 2168
891 uAsnArgSerArgSerLeuSerAsnSerAsnProAspIleSerGlyThrP 908
2169 ..... ACAAAAGTGTGAAGAACTAC..... GTGGACGCT... 2198
908 roThrSerProAspAspGluValArgSerIleIleGlySerLysGlyLeu 924
2198 ..... 2198
925 AspArgSerAsnSerTrpValAsnThrGlyGlyProLysAlaIleProTr 941
2198 ..... 2198
941 pGlySerAsnProSerProSerAlaGluSerThrGlnAlaMetAspArgS 958
2199 .GCTGAGAACCCCGGAGATAAATAGACAGCTGTGTACAAGCATG..... 2240

```

```

958 erCysasnArgMetSerSerHisThrGluThrSerSerPheLeuGlnThr 974
2240 .....
975 LeuThrGlyArgLeuProThrLysLysLeuPheHisGluLeuAlaLe 991
2240 .....
2240 .....
991 uGlnTrpValValCysSerGlySerValArgGluSerAlaLeuGlnGln 1008
2241 .....
1008 IaTrpPhePheGluLeuMetValLysSerMetVal.....HisHis 1022
2256 .....ATCTTCAAGTTCATCGTCGCGCTCCAGATCTCTGTTCAATCA 2297
1023 LeuTyrrPheAsnAspLysLeuGlnAlaProArgLysSerArgPheProG 1039
2298 CTGTATGAAAC...AAGGAGAGAGCTGAC..... 2324
1039 uArgPheMetAspPheAlaAlaLeuValSerThrIleAlaSerAspI 1056
2325 .....TTCTGGAACTCTTCTGCTCCAGCTCTTCAGTCCATC...AATGACA 2367
1056 LeuAlSerArgPheGlnLysAspThrGluMetValGluArgLeuAsnThr 1072
2368 TGATGAGCAGCATGTCCAGACCGACCGTCCGGGTGAAGGGGACACACTG 2417
1073 Ser...LeuAlaPhePheLeuAsnAspLeuLeuSerValMetAspArg 1088
2418 AATACTTACCAACGATCGTCAACGATGGAATGTTGTTGAT..... 2462
1088 yPheValPheSerLeuIleLysSerCysTyrlLysGlnValSerSerLys 1105
2462 .....
1105 euTyrrSerLeuProAsnProSerValLeuValSerLeuArgLeuAspPhe 1121
2463 .....CCCAAGAGCTCAGCAAAATGTTTACTGTAATTC 2495
1122 LeuArgIleIleCysSerHisGluHisTyrrValThrLeuAsnLeuProC 1138
2496 ATC.....CTCAATGTTCCCAT 2512
1138 sSerLeuLeuThrProProAlaSerProSerProSerValSerSerAlaT 1155
2513 GGCGCTTGTGACC..... 2525
1155 hrSerGlnSerSerGlyPheSerThrAsnValGlnAspGlnLysIleAla 1171
2526 .....ATCCAGAAACTTAC 2540
1172 AsnMetPheGluLeu.....SerValProPheArgGln.....G 1183
2541 TGCTTGATCGAATGTCACAGTGCACCTCTTACACAGCATGACATGACG 2590
1183 nhIstTyrrLeuAlaGlyLeuValLeuThrGluLeuAlaValIleLeuAsp 1199
2591 AGAGATCCGCTTCCATGATGACCGATCAGCTCAAGTACCATCTGGAGA 2640
1200 .....ProAspIleGluGlyLeuPheGlyLeuHisLysLysValIleAsn 1214
2641 GACAGAGAGACCTGGAGCGCTGTCAGCTGCTCAGCCACATCCCTGGAG 2690
1215 MetValHisAsnLeuLeuSerSerHisAspSerAspProArgTyrrSerAs 1231
2691 GTGCTGTACAGG.....AAGGAGGTGGGCCCAACCCAGAGGCA 2728
1231 pProGlnIleLysAlaArgValAlaMetLeuTyrrLeuProLeuIleGlyI 1248
2729 CGTCAG.....A 2736
1248 IeIleMetGluThrValProGlnLeuTyrrAspPheThrGluThrHisAsn 1264
2737 TTATCATGAGAGAACTT.....CTCCGACCGGTGAACGACCGCATT 2780
1265 GlnArgGlyArgProIleCysIleAlaThrAspAspTyrrGluSerGlu 1281
2781 TCCATGGGACGA.....GATTCGAA... 2801
1281 rGlySerMetIleSerGlnThrValAlaMetAlaIleAlaGlyThrSerV 1298
2802 .....CTATTGGAACCTTCGGCTTCATGACACT.....ATT 2838
1298 aLProGlnLeuThrArgProGlySerPheLeuLeuThrSerThrSerGly 1314
2839 TACACAAATGGAAGATTACCATTTATGCCCATTTGATCAAGACTTTGGG 2888
1315 ArgGlnHisThr.....ThrPheSerAlaGluSer.....Se 1325
2889 AAATGAGAGACTGATGTGTAGATTTCCTAATGAAACATTCATCATGTT 2938
1325 rArgSerLeuLeu.....IleCysLeuLeuTrpValLeuLysA 1338
2939 TAAGAACCTCATTTGGAAGAAAGACTTTACCCCTTCGACTGGTGATCATGA 2988
1338 sn.....AlaAsp 1340
2989 ACATGGTGCAAATAAAGTCTCTCTCGAGACCATTTAATCATGATGACAT 3038
1341 GluThrValLeuGlnLysTyrrPheThrAspLeuSerValLeuGlnLeuAs 1357
3039 .....ATGCTGAACAATAAATTTCTGATCAAGCCAACTTGATGACTACA 3082
1357 nArgLeuLeuAspLeuLeuTyrrLeuCysValSerCysPheGluTyrrLys 1374
3083 GCTGTGAACACTACTTCTCCTGCTGCT..... 3116
1374 LysLysValPheGluArgMetAsnSerLeuThrPheLys...Ser 1389
3117 .....TTCTTACTACAGAGTCCCTGCACAGGAGAAATTTTCA 3155
1390 LysAspMetArgAlaLysLeuGluGluAlaIleLeuGlySerIleGlyAl 1406
3156 AGTCCCAAGAGGCCAA.....ATCCTTAACAAGTACGCA... 3191
1406 aArgGlnGluMetValArgArgSerArgGlyGlnLeuGluArgSerPro 1423
3192 .....GATATGAGAGACAGATTGGCTTGAATC..... 3221
1423 erGlySerAlaPheGlySerGlnGluAsnLeuArgTrpArgLysAspMet 1439
3222 .....ACAGACATG 3230
1440 ThrHisTrpArgGlnAsnThrGluLysLeuAspLysSerArgAlaGluI 1456
3231 TGTATC..... 3236
1456 eGluHisGluAlaLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeu 1473
3237 .....AACCTTGTCAACACAGATAAAGT 3261
1473 IeIleLeuAspThrLeuGluIleValAlaGlnThrValSerValThrGlu 1489
3262 TCATTCCGAATAGTGGGCCCAATATTAGAAATGACATTAAATTCGCCAG 3311
1490 SerLysGluSerIleLeuGlyGlyValLeuLysValLeuHisSerMet 1506
3312 ACGGAG.....CTGCCAAAGCCACCATCCTCTTCTTGTATGAT 3355
1506 tAlaCysAsnGlnSerAlaValTyrrLeuGlnHisCysPheAlaThrGln 1523
3356 GCAAGTGAATTCATTCCATTCGACCCGAGCTTCCAATGTTT...GAAA 3399
1523 rGAlaLeuValSerLysPhePro..... 1530

```


2076 ValGlnLysGluTyr 2080
 |||:|||||
 4845 GTGGAGAAAGACTAC 4859

seq_name: /cgn2_6/ptodata/2/ina/bb_comb.seq:us-09-251-645-11

seq_documentation_block:

; Sequence 11, Application US/09251645

; Patent No. 6281413

; GENERAL INFORMATION:

; APPLICANT: Kramer, Vance C.

; APPLICANT: Morgan, Michael K.

; APPLICANT: Anderson, Arne R.

; APPLICANT: Hart, Hope

; APPLICANT: Warren, Gregory W.

; APPLICANT: Dunn, Martha

; APPLICANT: Chen, Jeng S.

; TITLE OF INVENTION: NOVEL INSECTICIDAL TOXINS FROM PHOTORHABDUS LUMINESCENS

; TITLE OF INVENTION: AND NUCLEIC ACID SEQUENCES CODING THEREFOR

; FILE REFERENCE: CGC1963/A

; CURRENT APPLICATION NUMBER: US/09/251,645

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 11

; LENGTH: 37948

; TYPE: DNA

; ORGANISM: Photorhabdus luminescens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (15171)..(18035)

; OTHER INFORMATION: orf5

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (23768)..(31336)

; OTHER INFORMATION: hph2

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (31393)..(35838)

; OTHER INFORMATION: orf2

; US-09-251-645-11

alignment_scores:

Quality: 193.50 Length: 2014

Ratio: 0.237 Gaps: 94

Percent Similarity: 40.616 Percent Identity: 17.676

alignment_block:

US-09-737-246-2 x US-09-251-645-11 ..

Align seg 1/1 to: US-09-251-645-11 from: 1 to: 37948

38 AsnIleValGlyAsnIleSerHisHisThrValProLeuThrGluAl 54

||| :|||:||||| :||| :|||

4243 AATGATATAGAGATCTTCTCATCTCTAGACAAATTTACTGGCAGA 4292

54 aValAspProValAspLeuGluAspTyrLeuIleThrHisProLeuAlaVal 71

||| :||| :||| :|||

4293 TATC.....CATCAGTTAACA 4309

71 AlaSerGlyProLeu.....Argasp 78

||| :||| :||| :|||

4310 TTGATGAACGGGGTTACTGTTGGAACCATAGTAAGAAACCAACAT 4359

79 LeuIleGluPheProProAspAspIleGluValValTyrSerProArgAs 95

||| :||| :||| :|||

4360 TTATTCGAGATTACCTCGACATCTGTT..... 4389

95 pCysArgThrLeuValSerAlaValProGluGluSerGluMetAspProH 112

||| :||| :||| :|||

4390ACTCTAATTAAACAACCTTATGCT..... 4413

112 IsValArgAspCysIleArgSerTyrThrGluAspTrrPalAla..... 126

||| :||| :||| :|||

4414 ..GTCACTACTGCTACGT.....ACACAAAAGGAGAGTGTCTACG 4455

127ValIleArgLysTyrHisLysLeuGluThrGlyPheAs 139

||| :||| :||| :|||

4456 TTGTTTATGATGACTACTGATAATATAACAAAACCTTAACCCCGGAAT 4505

139 nProAsnThrLeuAspLysGlnIlysgIuArgGlnIlysgIuProLysG 156

||| :||| :||| :|||

4506 ACGGAATTACTGATACCGCTTACCAATGGCTTCAAGATTTTAATTAAGA 4555

156 lValPheGluSerAspGluAlaProAspGlyAsnSerTyrGlnAspAsp 172

||| :||| :||| :|||

4556 AGATGTGAAGCTGAAGAA.....GAT 4578

173 GlnAspAspLeuLysArgArgSerMetSerIleAspAspThrProArgG 189

||| :||| :||| :|||

4579 CTAGAGAAACCAAAAAGAAATTCAGAGCCGCAAGAAAT..... 4620

189 ySerTrrPalAcysSerIlePheAspLeuLysAsnSerLeuProAspAla 206

||| :||| :||| :|||

4620 4620

206 euleuProAsnLeuLeuAspArgThrProAsnGluGluIleAspArgGln 222

||| :||| :||| :|||

4621CTGGAATAATTCGGGAAACCAAGCCACACACCTCA 4656

223 AsnAspAspGlnArgLysSerAsnArgHisLysGluLeuPheAlaLeuH 239

||| :||| :||| :|||

4657 GAAGACAGAAACCAAGCCAGACAGACCTGAATAAAGCTGMAAGACCCA 4706

239 sProSerProAspGluGluProIleGluArgLeuSerValPro.... 254

||| :||| :||| :|||

4707 CGAAAAGCCGAGAAAACCTAGATGAGTCAGGAAAAAATCTCCCAAG 4756

255AspIleProLysGluHisPhe 261

||| :||| :||| :|||

4757 CCATATCTCCTTATATGCGCCGCTGCAATTACCATGTGAACATGCG 4806

262 GlyIleArgLeuLeuValLysCysLeuSerLeu..... 272

||| :||| :||| :|||

4807 GCATATTCATATCTCATCTGGCGAGATATCTGMAACCCGCGATAGGAA 4856

272 272

4857 AATGACAGCGGAAAAATTATGAACTGTTGCGGAAAAAATCCGTTACGG 4906

273LysPheGluIleGluIleGluProIle..... 281

||| :||| :||| :|||

4907 CTCAACCTGGAATTCACAAAACAACTGACCTGTGGTCCAGTATGGCAG 4956

282 ..PheAlaSerLeuAlaLeuTyrAspValLysGluLysLysLysIleSe 297

||| :||| :||| :|||

4957 CGCCTCGCAACATAGCGTTG.....ATTACCGTCTTACCGGCTTAA 5000

297 rGluAsnPheTyrPheAspLeuAsnSerGluIleMetLysGlyLeuLeuA 314

||| :||| :||| :|||

5001 CGAAAACACCTTAAGCTGTTGTGACAAAAGCCGCAACCTTGTGTATTA 5050

314 rGProHisValProProAlaIleThrThrLeuAlaIleArgSerAlaIle 330

||| :||| :||| :|||

5051 AAACCAAGCAACCCGAAACCAACCTGAACCA..... 5082

331 PheSerIleThrTyrProSerGlnAspValPheLeuValIleLysLeuG 347

||| :||| :||| :|||

5083CCAGCACATGACGTATCAACACACTATGTCACATAAC 5117

347 uLysValLeuGln...GlnGlyAspIleGlyGluCysAlaGluProTyrM 363

||| :||| :||| :|||

5118 GCGTTTACTGACGTGGGTTAAGCTAGAGGAAACCGCTTCTGTGAC 5167

```

363 etlle...phelysgluAlaaspaIaThrLys..... 372
   :: |||||:::
5168 TAACCAATTTAAAGAGACATTAACAGCAACTATGGCTAAGCT 5217
373 .....AsnLysGluLysLeuLysLeuLysSerGlnAlaasp1 386
   ::|||::: |||||::: |||||
5218 ATGAATCTTGATAAAATCTACTGGAGCAAGCAAGATTCAGCAAAAAC 5267
386 n..PhecyGlnArgLeuLysIlystyrArgMetProPheAla.TripIra 402
   ||||| ||||| ||||| ||||| ||||| |||||
5268 TGATTTTCCAA...CTGGCAATCTATGACAACTATGTCAGTGGATTA 5314
402 1aIleHisLeuMetAsnIleValSerSerAlaGlySerLeuGluIuArgsp 418
   ||| :::: ||||| |||
5315 ACATCTCAGTCATTAAGACATCTCTCCCAAGCATT..... 5352
419 SerThrGluValGluIleSerThrGluIuArgLysGlySerTrpSerG1 435
   ||||| ::::: |||||::: |||||::: |||||:::
5353 TCCCACTGACTCAAGTATTTGACCCAGAACCTCCCGTAACATAACCCA 5402
435 uArgArgAsnSerSerIleValGlyArgArgSerLeuGluIuArgThrTrs 452
   ::|||::: |||||::: |||||::: |||||:::
5403 ATGGGAAAGCGCGCTGCATATTAACGCGCGGCGAGACACCAAAAGA 5452
452 erGlyAspaSpAlaCysAsnLeuThrSerPheArgProAlaThrLeuThr 468
   ::::: ||| :::: ||| |||||::: |||||:::
5453 CTACGCCCTACATGGTTCTGATGAGTCCTCGAGTGTCTGCTTAAAGC 5502
469 ValThrAsnPhePheLys...GlnGluGlyAspArg..... 479
   ::::: |||||::: |||||
5503 ACATACTATATTTATTCATATCAACCAAAAGATCGAAGCAAGCAAAACA 5552
480 .....LeuSerAspGluAspLeuThrLysPheLeuAlaaspMetA 493
   |||||::: |||||::: |||||::: |||||:::
5553 TACAGTAATTAATAAACCGTATGATATATCAATATCCTATTGATCGATA 5602
493 rArgProSerSerValLeuArg.....ArgLeuArgProIleThrAla 507
   ::::: |||||::: |||||::: |||||
5603 ACCAAGTTCCCGCGCATTAATAACTACAGAGATCGTAGAGCTATCGCT 5652
508 GlnLeuLysIleAspIleSerProAlaProGlu..... 518
   ::|||::: |||||::: |||||
5653 AGTATCCAACTGATATTAACCGCGCTGAATAAATATGAGGAGATAC 5702
519 ..... 519
5703 TGTCAAGATGTCACCAAGCGCTCTTCTCACCACTGGGATAATACA 5752
519 snProHisTyrCysLeuThrProGluLeuGlnValLysLeuTyrPro 535
   ||::|::| |||||::: |||||
5753 ATAAACGCTACAGCACTTGGCGGCTATGGCTAAACCTCTTACTATTCGA 5802
536 AspSerArgValArgProThrArgGluIleLeuGluPheProAlaArgAs 552
   ::|||::: |||||::: |||||::: |||||
5803 GAGAAATACATGATCGACAGCTAGTATGGGACACAAATAATGATGGA 5852
552 pValTyrValProAsnThrThrTyrArgAsnLeu..... 563
   ||::: |||||::: |||||
5853 TACGTTGCTGCATTCATCAGCAAGCAATTAATATCGATACCGTAG 5902
563 ..... 563
5903 AAGATGCTTTAAATCTTACCTAACATCATTCGAAGAGGTGGTAATCTG 5952
564 .....LeuTyrIleTyrProGlnSerLeuAsnPheAlaAsnArgGlnG1 578
   ||||| ||||| ||||| ||||| ||||| |||||
5953 GAATCTCAGCGCTACCAACATCAATAT.....AATGATCAAGG 5996
578 Y.....SerAlaArgAsnIleThrValLysValGlnPheMetY 591
   |||||::: |||||::: |||||
5997 ATTAACCTACTTTATCGGACGTAGTAACAGAACTGATCAATATTTATT 6046
591 Yr.....GlyGluAspProSerAsn 597

```

```

6047 GGGCCAGTGTGGATTCACATTAATTCAGCGAAGGTAAATTCGCCGTAAT 6096
   :: |||||::: |||||::: |||||
598 AlameT..... 599
6097 GCGTCGAGTGAGTGTCACAAATTCATGATTCATTAACCCCTAACAGA 6146
600 .....ProValIlePhe..... 603
   ||||| ||||| ||||| ||||| |||||
6147 TACTATTACCCGGTAATTTTCCAAATCTCCGTATATCTGTGCTGG 6196
604 .....GlyLysSerCysSerGluPheSerLysGluAla..... 615
   |||||::: |||||::: |||||::: |||||:::
6197 AACAAAAAAGCGCCTTAACAGAGAGTATAAACCGCTGGGTTAT 6246
616 ...TyrThrAlaValIleTyrHisAsnArg.....Se 625
   ||||| ||||| ||||| ||||| |||||
6247 TATTATGAACTAAGAAATTAAGCGATATCCGTTATGACGCGACCTGGAATAC 6296
625 rPro.....AspPheHisGluGluIleLys.....V 634
   ::|||::: |||||::: |||||::: |||||:::
6297 ACCAGTCACCTTTGATGTATTAACCAAAATATCCGATTTAATCTGGAA 6346
634 allYsLeuPro.....AlaThrLeuThrAspHisHisLysLeu 646
   ||||| ||||| ||||| ||||| |||||
6347 ATAAACACCTCGACTTACTGCTCAAGCTTTCACAGCAGAGATGATTC 6396
647 LeuPheThrPheTyrHisValSerCysGlnGlnLysGlnAsnThr..... 661
   ||||| ||||| ||||| ||||| |||||
6397 CTGCTGATGTTTATTAATAAACAAGATCAATTAATCAATACACAAAC 6446
662 ...ProLeuGlu..... 664
   |||||::: |||||::: |||||::: |||||:::
6447 AGTACCAATTAAGAGACTATATATCACTCCAAATATGCTTTAAGGAA 6496
665 ...ThrProVal..... 667
   ||||| ||||| ||||| ||||| |||||
6497 TGAACCTGAAATTCACAAACCTAACGCTATTAACAGTTTGATTAAT 6546
667 ..... 667
6547 AGTATATTTGTCATATAATCCGTATGCAAGAACTAGCAAAATCCCTTC 6596
668 .....GlyTyrThrP..... 671
   ||||| ||||| ||||| ||||| |||||
6597 ATCAGTAATATGTAATACGTTATATGATGAGATGCGTATCTGAGTA 6646
671 ..... 671
6647 TGGTGTATGGCGGAATTTTCAGCCATCAACTGAGTCCATCATCAGAT 6696
672 .....IleProMetLeuGlnAsnGly..... 678
   ||||| ||||| ||||| ||||| |||||
6697 AAGTTAAACTCTACCAAGGTAAAGAAATATTCATTAATGAGACTTGA 6746
678 ..... 678
6747 CCGACAAAGCAACCAATGCAACTGATGAAGAATACGCTGACCTGGTG 6796
679 .....ArgLeuLys 681
   ||||| ||||| ||||| ||||| |||||
6797 ATAAATTTATTTATTTACTACTAGTATTAACCCCAATATTTGTG 6846
682 ThrGlyGlnPheCysLeuProVal..... 689
   ||||| ||||| ||||| ||||| |||||
6847 AATAAAAAATTCATTTACCTGTTTATCAGTATAGTGAACACTACCAA 6896
689 ..... 689
6897 TAATGAGAAAGACGCTGCTGTTTATGAGAAAGTACTACTACTTTCG 6946
690 .....SerLeuGluLysProProGlnAlaTyrSerVal 700
   ||||| ||||| ||||| ||||| |||||

```

```

6947 TAAGAGCCTGGTTCCTTAACCTTCCTGCTGCTCAAGAAATGCCACA 6996
701 Leu..... 701
6997 ACCACTGGCGGATAGTGTACTATGTATATTGATAACAAACA 7046
702 .SerProGluValProLeuProGlyMetLysTrpValAspAsnHisLysG 718
    ||| :|||:||||| :|||:|||||:|
7047 TAGTGACGATGTTCATTAACAATATTTCATATGTGATGACACGGGTG 7096
718 LysAlaPheAsnValGluValAlaValSerSerLeuHisThrGlnAsp 734
    |||: :|||:||||| :|||:|||||:|
7097 GATTT...GACACGTGATGTTTCAGGGATATTATCTATTAAATGACGACATT 7143
735 ProTrpLeuAspLysPhePheAlaLeuValAsnAla.....LeuAspGlu 749
    :|||: :|||:||||| :|||:|||||:|
7144 AATCAT...TCAAAAGTTAAAGTAAATGATGAAAGCCGAAAGTATCACGA 7190
749 uHisLeuPhe.....ProValA 755
    |||: :|||:||||| :|||:|||||:|
7191 GCAAACTTTTGTAGCGAGAAACAGTAATGTCCCAACCAATCCGTCGC 7240
755 rGllleGlyAspMetArgIleMetGluAsnAsnLeuGluAsnGluLeuLys 771
    |||: :|||:||||| :|||:|||||:|
7241 GCTTGAGAAATGAAATGATTCAGTTTAAAGAGCTTGAATAGATATCTCC 7290
772 SerSerIleSerAlaLeuAsnSerSerGlnLeuGluProValValArgPh 788
    :|||: :|||:||||| :|||:|||||:|
7291 ACACGTGACATTTTCATTAATTAATGACCAAGATTTGAT.....ATCACCTT 7334
788 eLeu.HisLeuLeuLeuAspLysLeuIleLeuLeuValIleArgProPro 804
    |||: :|||:||||| :|||:|||||:|
7335 TATCGCATTTGCTGAGAAATTTGACGATATAATAGATGATCGT..... 7377
805 ValIleAlaGlyGlnIleValAsnLeuGlyGlnAlaSerPheGluAlaMe 821
    |||: :|||:||||| :|||:|||||:|
7378 .....AACTTGGCGCAAGAACATTTTCAGT..... 7401
821 tAlaSerIleIleAsnArg.....LeuHisL 830
    |||: :|||:||||| :|||:|||||:|
7402 .ATTGCTATTATCAAAAAAGCGAAACTGATATGCTCCCTGACCCCTGCACC 7450
830 yAsnLeuGluGlyAsnHisAspGlnHisGlyArgAsnSerLeuLeuAla 846
    ||| :|||:||||| :|||:|||||:|
7451 AATAAGCAAAAGCGGCGCAATATATGACAGTGGGAAACTCTTGATTT... 7497
847 SerTrpIleHisTrpValPheArgLeuProAsnThrTrpProAsnSerSe 863
    |||: :|||:||||| :|||:|||||:|
7498 .....CGCCTT...AATACG..... 7509
863 rSerProGlyProGlyGlyLeuGlyGlySerValHisTrpAlaThrMeta 880
    :|||: :|||:||||| :|||:|||||:|
7510 .....CTATTGTG 7516
880 laArgSerAlaValArgProAlaSerLeuAsnLeuAsnArgSerArgSer 896
    |||: :|||:||||| :|||:|||||:|
7517 CCCGTCATTAATATAGCCGACCAAGCGGGGATAGATATATTTTGGAC 7566
897 LeuSerAsnSerAsn..... 901
    :|||: :|||:||||| :|||:|||||:|
7567 ATGGACACTCAGAAATTTTCAGAACCTAAATTAAGAGAAATTTCTCGTA 7616
902 .....ProAspIleSerGlyThrProThr..... 909
    |||: :|||:||||| :|||:|||||:|
7617 TGCCTATGACCAATGAGACTTCAACGGCGCCCAAGCCTATATTTCTGGG 7666
910 .....SerProAspAspGluValArgSerIleIleGlySer 921
    :|||: :|||:||||| :|||:|||||:|
7667 AACTGTTCTACTACAAACCCGATGCTGATTTGCTCAACGTTTGTCTGACAGAA 7716
922 LysGlyLeuAspArgSerAsnSerTrpVal.....AsnThrGln 934
    :|||: :|||:||||| :|||:|||||:|
7717 CAAAACTTCGATGAGGCTAACCGTTGGCTGAATATGTCTGGAACCCATC 7766

```

```

934 uGly..... 935
    :|||:
7767 CGGTATATTTGTCATATGTCATTAATGCAACATTCACCCGTGAATGTTCCGC 7816
936 ....ProLysAlaAlaProTrpGlySerAsnProSerProSerAlaGlu 950
    :|||: :|||:||||| :|||:|||||:|
7817 CATTAACAAGAACACTAGTTTGAACGATGATCCGTGATCATTTGAT 7866
951 SerThrGlnAlaMetAspArgSerCysAsnArgMetSerSerHisThrGln 967
    |||: :|||:||||| :|||:|||||:|
7867 CCT.....GATACCATAGCTCAACATGATCCAAATGCACTACAGA 7904
967 uThrSerSerPheLeuGlnThrLeu.....ThrGlyArgL 979
    :|||: :|||:||||| :|||:|||||:|
7905 AGTCGCCACCTTTATGCGCACCCCTAGATCTGTGATGCAAGGAGGAGATT 7954
979 euProThrLysLysLeuPheHisGluGluLeuAlaLeuGlnTrpValVal 995
    :|||: :|||:||||| :|||:|||||:|
7955 ACGCCTATGCGCAATTTGAGCGGGACACTCGCT..... 7989
996 CysSerGlySerValArgGluSerAlaLeuGlnGlnAlaTrpPhePheP 1012
    |||: :|||:||||| :|||:|||||:|
7990 .....GAAGCCAAATG.....TGGTATATGCA 8012
1012 eGluLeuMetValLysSerMetValHisHisLeuTrpPheAsnAspLysL 1029
    |||: :|||:||||| :|||:|||||:|
8013 GGCACCTG.....CATCTA...TTGGTGATTAAC 8038
1029 euGluAlaProArgLysSerArgPheProGluArgPheMetAspAspIle 1045
    :|||: :|||:||||| :|||:|||||:|
8039 CTCATCTACACACCTCAGTTTCAGCATGATGATCCAGAGCTAGAAAGGCC 8088
1046 AlaAlaLeuValSerThrIleAla.....SerAspIleVal 1057
    |||: :|||:||||| :|||:|||||:|
8089 GCACCTCTTCAAAAACACAGCCACATGCCCAAGAAATACACAGATTTCAG 8138
1057 LysArgPheGlnLysAspThrGlnMetValGluArgLeuAsnThrSerL 1074
    :|||: :|||:||||| :|||:|||||:|
8139 ACAAGGACTTCCTACATCCACAGGGGTCTAAAGATGAATGAAACAGATC 8188
1074 euAlaPhePheLeuAsnAspLeuLeuSerValMetAspArgGlyPheVal 1090
    |||: :|||:||||| :|||:|||||:|
8189 TT.....TTCCTGCCCAAGTCAACGAACTGATG..... 8217
1091 PheSerLeuIleLysSerCysTrpLysGlnValSerSerLysLeuTrpSe 1107
    |||: :|||:||||| :|||:|||||:|
8218 .....CTGACCTACTGCGCAGAAACCTAGAACACAGGTTGTATTA 8255
1107 rLeuProAsnProSerValLeuValSerLeuArgLeuAspPheLeuArgI 1124
    :|||: :|||:||||| :|||:|||||:|
8256 CCTGGGCCAT.....AACCTCTCTATTGAT..... 8280
1124 LeileCysSerHisGlnHisTrpValThrLeuAsnLeuProCysSerLeu 1140
    |||: :|||:||||| :|||:|||||:|
8281 .....GTCACACCTTTACATTTGGCT.....ATP 8304
1141 LeuThrProProAlaSerProSerProSerValSerSerAlaThrSerGln 1157
    :|||: :|||:||||| :|||:|||||:|
8305 TTCGCAACACCCAGCAGATGCCAAAGCCCTGCTCAGGCGCCCTGTCGCA 8354
1157 nSerSerGlyPheSerThrAsnValGlnAspGlnLysIleAla..... 1171
    :|||: :|||:||||| :|||:|||||:|
8355 TTCACAAAGGT...GGAAAGTATCTTCATCATGAAATTTATATTCAGTGTGC 8401
1172 .....AsnMetPheGlu..... 1175
    :|||: :|||:||||| :|||:|||||:|
8402 GTTCCCTCATATATGCTGAAAAAGCCCGCTAGTATGTCATCAGCTAAC 8451
1176 .....LeuSerValProPheArgGln..... 1182
    :|||: :|||:||||| :|||:|||||:|
8452 CAATTGCGTCCACATTTGCAAAATATATATGCAACGTCACAGATCGGAGGC 8501

```

1183GlnHisTyrLeuAlaGlyLeuValLeuThrGlnL 1194
12222
8502 ATTAAACAGCTGTGTGCAAAATACAGCGGAGAACTGATATGACCAATC 8551
1194 euAlaVal.....IleLeuAspProAlaGluGlyLeuPhe 1206
1194
8552 TCAGCATACAGACAAACCAATCTCAGATGAAGACATCAACGGGGTGAAA 8592
1207 GlyLeuHisLysLysValIleAsnMetValHis...AsnLeuLeuSerSe 1222
1194
8593AAACCTGTCTAGAAAAAACCAGCGGAGACCCAGCTCCG 8633
1222 rHisAspSerAspProArgTyrSerAspProGlnIleLysAla..... 1236
1194
8634 TTTGATAGCTACAGCAAAATCTCAGATGAAGACATCAACGGGGTGAAA 8683
1237ArgValAlaMet 1240
8684 AACAGCAATGCGGTGCGTCTCCGCTGCGCATCTCAACGCCCTT 8733
1241 LeuTyrLeuProLeuIleGlyIleLeuMetGluThrValProGlnLeuTyr 1257
8734 CAGACATCATCTGTGCGGCGGAGCACTTGAATCGCGCCCAACATCTT 8783
1257 rAspPheThrGlnThrHisAsnGlnArgGlyArgProIleCysIleAlaT 1274
8784 CGCTTCGCTGATGGTGGCAGCGCTTGGGGG...GCGATCGCCCAAGCCA 8830
1274 hAspAspIrrGluSerGluSerGlySerMetIleSerGlnThrValAla 1290
8831 CAGGTAATCTCATGAGATTCCTCCGCACTGTTATGAC..... 8868
1291 MetAlaIleAlaGlyThrSerValProGlnLeuThrArgProGlySerPhe 1307
8869ACCAACCGCATAAATACACCAATCTGAAGCCTA 8903
1307 eLeuLeuThrSerThrSerGlyArgGlnHisThrThrPheSerAlaGln 1324
8904 C.....CGTCGCGCTGTCAGGAA..... 8922
1324 eSerArgSerLeuLeuIleCysLeuLeuTyr...ValLeuLysAsnAla 1339
8923TGGAAATTCACCGTAAATAC 8943
1340 AspGluThrValLeuGlnLysTrpPheThrAspLeuSerValLeuGlnLe 1356
8944 GCGGAGGCGAGACTGAACAATCATGCTCACTGTTGCGTGGCGAT 8993
1356 uAsnArgLeuLeuAspLeuLeuTyrLeuCysValSerCysPheGluTyrL 1373
8994 GCGCGGTGAAGCCGCGATATTG.....CAGAAA 9022
1373 yGlyLysLysValPheGluArgMetAsn.....SerLeuThrPheLys 1387
9023 CCAGCTTAAACCCACAAAGAGCACTCATGCAACAATCACTTCTCTG 9072
1388 LysSerLysAspMetArgAlaLysLeuGlnGluAlaIleLeuGlySerI 1404
9073 CAACATTAAGTTCATGATCAGCGCTGTACACATGCGCTGCGTGCATT 9122
1404 eGlyAla.....A 1407
9123 GTCGCCCATTTACTTCAGTTCTATGATTAAAGGTAGCTGCTGTTGA 9172
1407 rGlnGlnLysMetValArg.....SerArg... 1415
9173 TGGCGGAATAGCTATGCTGCGGAGACTAAAGATACCGCATCAGCGCTT 9222
1416GlyGlnLeuGlnLysSerProSerGlySerAlaPheGlySe 1429
9223 ATCAAAACCGCGGCTGCGCAGGGAACCATGCGCTTTCCTGCGGCGTGA 9272
1429 r.....GlnGluAsnLeuArgTyrPargLysAspMetThrHisTrpArg 1444

9273 AACCTATATCTGATCTGGCAGATGGAAGAT...GCCCACTGGAAC 9319
1444 LnsAsnThrGlnLysLysLeuAspLysSerArg.....AlaGluIle 1456
9320 AGATTAACGCGTACTGGAGTAGAACGTACCGTTTCGCTGGCGGAAGTC 9369
1457GlnHisGluAl 1460
9370 TATGCCAATTTACCGCAAGATTAATTTATCTGACTCAGCAAAATTAAGAA 9419
1460 aleuIle.....AspGlyAsnLeuAlaThrGluAlaAsnLeuIle 1474
9420 GTTGAGTAGTAAGGTTCAGCAGCGCGGCGGAGCAAGCAATTAATTAAGCTG 9469
1474 leLeuAspThrLeuGlnIleValAlaGlnThrValSerValThrGluSer 1490
9470 CGTTTATACCAATACCAATACCTCTGTAGAACGTCATTTGTTATCT 9519
1491 LysGluSerIleLeuGlnGly...ValLeuLysValLeuLeuHisSerMet 1507
9520 ACCTTGACATTTAGCAGCGATTTCTGATTTGTTATGTTAAACCGCTG 9569
1507 lAcysAsnGlnSerAlaValTyrLeuGlnHisCys 1518
9570 TATTAAACAGATCAGCGTTACCTCCGCGCAGCACTGC 9604
seq_name: /cgn2.6/plodata/2/lna/6B_COMB.seq:US-08-952-127-11
seq_documentation_block:
; Sequence 11, Application US/08952127
; Patent No. 6211336
; GENERAL INFORMATION:
; APPLICANT: Shiloh, Yosef
; APPLICANT: Tagle, Danilo A.
; APPLICANT: Collins, Francis S.
; TITLE OF INVENTION: ATRAXIA-TELANGIECTASIA GENE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 6211336thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/952,127
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,995
; REFERENCE/DOCKET NUMBER: 2290,00029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 810-539-5055
; TELEFAX: 810-539-5050
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9620 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: Chromosome 9, Band 9C
; US-08-952-127-11

[illegible]


```

1066 eValaGlunArgLeuAsnThrSerLeuAlaPhePheLeuAsnAspLeu 1082
      |||||
5637 .....ACTCTGCTG 5645
1083 SerValMetAspArgGlyPheValPheSerLeuIleLysSerCysTyrLy 1099
      |||||
5646 TCTGGCAGCTCCGAGGATTT.....TTCACCTAGCTGTTTAA 5683
1099 sGlnValSerSerLysLeuTyrSerLeuProAsnProSerValLeuValS 1116
      |||||
5684 GCAT...TCTCCCAAGCAAGCCGCTCAGCAACTCTCTCAAAATTCGATT 5730
1116 eLeuArgLeuAspPheLeuArgIleIleCysSerHisGluHisTyrVal 1132
      |||||
5731 CAGAGTCAGAGAACTTCTCCGATGCTGTTGGATATAAAAGTCACAAAGA 5780
1133 ThrLeuAsnLeuProCysSerLeuLeuThrProProAlaSerProSerPr 1149
      |||||
5781 ACCATGCTTGTCTGCTGATCTGAGAAAGCAAAAGACCT..... 5825
1149 oSerValSerSerAlaThrSerGlnSerGlyPheSerThrAsnValG 1166
      |||||
5826 .....TCTCGGGACAGCTTTTGATGACGCTTTCTGGCTGGATTGA 5868
1166 InAspGlnLysIleAlaAsnMetPheGluLeuSerValProPheArgGln 1182
      |||||
5869 ATTATCTTAGGTTGGAGAGTGCGCTCACTCTGC.....TCT 5906
1183 GlnHisTyrLeuAlaGlyLeuValLeuThrGluLeuAlaValIleLeuAs 1199
      |||||
5907 GCACACTTCACGACC...TTGCTACGACGAGATCTAT..... 5942
1199 pProAspAlaGluGlyLeuPheGlyLeuHisLysLysValIleAsnMetV 1216
5942 ..... 5942
1216 alHisAsnLeuLeuSerSerHisAspSerAspProArgTyrSerAspPro 1232
      |||||
5943 .....TCAGATTAAGAAAGCACAGCAGG 5966
1233 GlnIleLysAlaArgValAlaMetLeuTyrLeuProLeuIleGlyIleI 1249
      |||||
5967 CAAGAGAAAAGAAATCCA..... 5984
1249 eMetGluThrValProGlnLeuTyrAspPheThrGluThrHisAsnGlnA 1266
5985 .....ACATTTGAGAGAGAGTC 6003
1266 rGlyLysArgProIleCysIleAlaThrAspAspTyrGluSerGly 1282
      |||||
6004 AAGGACAACTATTCTAGTTGAGTGAATAAAAGTAAGAGAACTGGA 6053
1283 SerMetLysSerGlnThrValAlaMetAlaIle..... 1293
      |||||
6054 ATAAAGCTTA...CAGGATCTTCTCTAGAGATCTACAGAAATATAGAGA 6100
1294 .....AlaGlyThrSerValProGlnLeuTyr 1302
6101 GCCGGACAGCCTGTATGCTGTGGAGAGGAAATGTTACAAACCCCTTA 6150
1302 hrArgProGlySerPhe.....LeuLeu 1309
      |||||
6151 CTGAGATACGACATATGAACATGAAGCTAGGAGAAAGCCTTAGTA 6200
1310 Thr.....SerThrSerGlyArgGlnHisThrTh 1319
      |||||
6201 ACTTAGACACCTGGAGACACACATCTCTCTCCACCCGACAGTCAGAGAT 6250
1319 rPheSerAlaGluSerSerArgSerLeuLeuIleCysLeuLeuThrVal 1336
      |||||
6251 CATCCAGGCCCTGCAGAAATTTGGGGCTCTCCCATATCCCTGTCTGTATTC 6300
1336 euLysAsnAlaAspGluThrValLeuGlnLysThrPheThrAspLeuSer 1352

```

```

6301 TGAAGGATTAGACTATGAA...AGACGAGAGATGTGCGCTGAGTCGACG 6347
      |||||
1353 ValLeuGlnLeuAsnArgLeuLeuAspLeuLeuTyr.....LeuCysVal 1367
      |||||
6348 GAGCTGGGTTTACAGGGGGGCGGAGAGACATGACGTGGGGCTGTGGCG 6397
1367 LserCys...PheGluTyrLysGlyLysLysValPheGluArgMet...A 1382
      |||||
6398 TTCTGCCGCCAAGAACTAGAGAGAACCACTTACATGAATCTGTATTA 6447
1382 snSerLeuThrPheLysLysSerLysAspMetArgAlaLysLeuGluGln 1398
      |||||
6448 ATGCTCTGCAAGTGTCTAAGAAACAGAGAAATTTCCACATTTTGAAGT 6497
1399 AlaIleLeuGlySerIleGlyAlaArgGlnMetValArgArgSerAr 1415
      |||||
6498 CTCGATATAGCCAGTCTTTTCAGGGTGAAAGAAAGTTGAAGTTGAGTA 6547
1415 gGlyGlnLeuGluArg.....SerProSerGlySer..... 1425
      |||||
6548 GGGCAGCCTTAGCTGTATATTCGCTGTAATCCACACTAGTAGATTGC 6597
1426 ..AlaPheGlySerGlnGluAsnLeu.....ArgTyrArgLysAspMet 1439
      |||||
6598 AGCAATTTGAGAACTGGAAACAGTGGCGAGCTTTCTCAAGTGCATGC 6647
1440 ThrHisThrPargLlnAsnThrGlnLysLeuAspLysSerArgAlaGlu 1456
      |||||
6648 ACA.....GACAGAGAGGCTCTGTGAGC 6670
1456 e.....GluHisGluAlaLeuIle...AspGlyAsnLeuA 1467
      |||||
6671 ATACTGGAGAGCGCAAGACACTCCGACTTCTGAAGAACAGCAGACTTCA 6720
1467 lathrGluAlaAsnLeuIleIleLeuAspThrLeuGluIleValValGln 1483
      |||||
6721 GCTTTCAGAGACCTCTCATGCTCTGCGCACA.....GTCAATCTGGAG 6764
1484 ThrValSerValThrGlnSerLysGluSerIleLeuGlyValLeuLys 1500
      |||||
6765 ACCCTG...GTACAGAAAGAAATGAGCGCTCTCAAGAGCATGCTCTAA 6811
1500 sValLeuLeu..... 1503
      |||||
6812 GGACATTTCTACCAAAACACCTGTTGAATTCCTGTTCTGGCTGAGACT 6861
1503 ..... 1503
6862 TCAGAACACACAGCTCCCTGAAGAGCAATATTCAAAATTAAACAAATAT 6911
1504 HisSerMetAlaCysAsnGlnSerAlaValLysLeuGlnHis..... 1517
      |||||
6912 AATTCAGCTATTGTGGAATTTGTAGTGGCATTTGGAAGAGACAAAGT 6961
1518 .....CysPheAlaThrGlnArgAlaL 1525
      |||||
6962 ATTCTGGGCAAAAAAGAGCAGAGCTTGTCTGAGATTTCTCAAGCAGA 7011
1525 euValSerLysPheProGluLeuLeuPheGluGluGluThrGluGlnCys 1541
      |||||
7012 TGATCAAGAGATTTGAGCTCCAGCTTAAAGATAAAGAAATGATGAGAGT 7061
1542 AlaAspLeu.....CysLeuArgLeuLeuArgHisCysSer 1554
      |||||
7062 CTCAAAGTCATATACGACAGCTGTCTGAGGGT.....TGTGGGAG 7102
1554 rSerIleGly...ThrIleArgSerHisProSerAlaSerLeuTyrLeuL 1570
      |||||
7103 CTGGGTGGCAGAAACTGTCTAGAAACCCGTCAGTCACTATGCAAGACT 7152
1570 euMetArgGlnAsnPheGluIleGlyAsnAsnPhe..... 1581
      |||||

```

```

7153 ATCTAGAAAAAGCGGTGAGAGTTCTGGAGATTACGATGGCAACAGCAGA 7202
1582 .....AlaArgValLysMetGlnValProMetSerLeuSerLeuVal 1596
       ::::: ::::: ::::: ::::: ::::: :::::
7203 GAGCTCGAAGATGGACACAGATGAAGCCCTTCTCTCGTTGGCAAGTTCTC 7252
1596 LGlyThrSerGlnAsnRheAsnGlnGluPheLeuArgArgSer..... 1610
       ::::: ::::: ::::: ::::: ::::: :::::
7253 TGATACTCAGACACAGAGATTGAAGAACTACATGAAGTCTCATCAATTTG 7302
1611 .....LeuLysThrIleLeuThrTyrAlaGluGluAspLeuGlu...Leu 1624
       ::::: ::::: ::::: ::::: ::::: :::::
7303 AAAACAGCAAACTCTCTTAAAGAGCCCAAGAGAGTGGCCCTTCTA 7352
1625 ArgGluThrThrPheProAspGlnValGlnAspLeuValPheAsnLeuH1 1641
       ::::: :::::
7353 AGGGAA.....CA 7360
1641 sMetIleLeuSerAspThrValLysMetLysGlnHisGlnGluAspProG 1658
       ::::: ::::: ::::: :::::
7361 TAAATTCAGACCCACAGATACACAGTAAGGTTCAGCGGAA..... 7403
1658 lMetLeuIleAspLeuMetTyrArgIleAlaLysGlyTyrGlnThrSer 1674
7403 ..... 7403
1675 ProGluArgLeuThrThrPheLeuGlnAsnMetAlaGlyLysHisSerGluAr 1691
7403 ..... 7403
1691 gSerAsnHisAlaGlnAlaAlaGlnCysLeuValHisSerAlaAlaLeuV 1708
       ::::: ::::: :::::
7404 .....CTGAGCTGAGACGATGTGCTCTCCG..... 7430
1708 alaIaGlyTyrLeuSerMetLeuGlnAspArgLys...TyrLeuProVal 1723
       ::::: ::::: :::::
7431 .....GCAGTGAAGAGATGCGCAAGCGCTTCTCG..... 7460
1724 GlyCysValThrPheGlnAsn..... 1730
       ::::: :::::
7461 ...TGTAAGCAGTGGAGAACTACATCAACTGCTTACTAAGCGGGAAGA 7507
1731 .....IleSerSerAsnValLeuGlnGluSerA 1740
7508 ACATGATCTGTGGGTGTCGCGCTTGTCGCCCTTGCGTTGAAGATTCTG 7557
1740 lValAlaSerAsp.....AspValAlaSerProAspGluGluGlyIleCys 1754
       ::::: ::::: :::::
7558 GAGTTTCTGAAGTCAATGGCATGATGAAGAAAGATGAAGATTGCA 7607
1755 SerGlyLysThrPheThrGlnSerGlyLeuValGlyLeuLeuGlnGlnAl 1771
       ::::: ::::: :::::
7608 TCTTAATAGTTT.....TTGCCCTCATGATCAAT 7639
1771 alaIaAlaSerPheSerMet.....AlaGlyMetTyrGlnAl 1783
       ::::: :::::
7640 GCGTGTCTGAATGGGCAAAATGACGGAGGCGCTAGATTTCACAGAG 7689
1783 lValAlaGlnGluValTyrLys..... 1789
       ::::: :::::
7690 TCTCTCAATATCTAATCTAGGATTCTACTGATCACCCTCATCATACT 7739
1790 .....ValLeuIleProIleHisGlnAlaAsnArgAsp..... 1800
       ::::: :::::
7740 TTGTTCATTATACTGCGCTTAGCAATGCGAACAAGATGATTTTTCAG 7789
1801 .....AlaLysL 1803
7790 CAATCAGAGACAAACAAGAGAGTCAATTAACCAAAAGTACATCTCAAG 7839
1803 yslLeuSerThrIleHisGlyLysLeuGlnGluAlaPheSerLysIleVal 1819
       ::::: :::::
7840 AAATCTCTACCCCTTGATGAGATCGAACAAGGCTGCACACAGAAATCATC 7889

```

```

1820 HisGlnSerThrGly.....TrpGluArgMe 1828
       ::::: :::::
7890 CACTCCATCAGAAATGAAGCATGATGAAGTGTGAAGACATGAGCGCGCT 7939
1828 tPheGlyThrThrPheArgValGly.PheTyrGlyThrLysPheGlyAsp 1844
       ::::: :::::
7940 CTGGCATGCGCTACATCATCTTGGCAACATGAGCCCTCTCAGTGAAGG 7969
1845 lLeuAspGlnGlnGluPheValTyrLys.GluProAla.IleThrLysLeu 1860
       ::::: :::::
7990 CTCAGAGAAAGCATCAATATATTCAGCCACACAGCCATCACTAAACTG 8039
1861 AlaGluIleSerHisArgLeuGlnGluPheTyrGlyGluArgPheGlyG1 1877
       :::::
8040 AAGAAATTAA.....GA 8050
1877 uAspValAlaGluValIleLysAspSerAsnProValAspLysCysLysL 1894
       :::::
8051 AGATGTTGTGTT.....CCCACTTGGAATTAAGG 8082
1894 euAspProAsnLysAlaTyrIleGlnIleThrTyrValGluProTyrPhe 1910
       :::::
8083 TTGATCCCAAGAGAGATGAAATCTGTGACTATTAATCAATTAA 8132
1911 AspThrTyrGlnMetLysAspArgIleThrTyrPheAspLysAsnTyrAs 1927
       :::::
8133 ACAGAAATTCGCTTACTGAGCGCTTA.....AA 8161
1927 nLeuArgArgPheMetTyrCysThrProPheThrLeuAspLysArgAlaH 1944
       :::::
8162 TTTACCCAAATTAATAGATTGTGTG.....GTTTGTATGCAAGGAA 8205
1944 lGlyGluLeu..... 1947
       :::::
8206 GGAGACAGCTTGTGTAAGGCGCGTGATGACTGAGCAAGATGCTGCATG 8255
1948 .....HisGluG1 1950
8256 CAGCAGTCTTCAGAGTGTCAATACACTACTGCAAGAAACACTGAGAC 8305
1950 nPheLysArgLysThrIleLeuThrThr.SerHisAlaPheProTyrIle 1966
       :::::
8306 TAGAAGAGGAAACTGATATCTGCACATACAGAGTGTGCCCTTCTC 8355
1967 LysThrArgValAsnValThrHisLysGluGlu..... 1977
       :::::
8356 AGCGAAGCGGTGTTCTCGAGTGTGCACAGGAAACGTTCTATTGTGAA 8405
1978 .IleIleLeuThrProIleGluValAlaIleGluAsp 1989
       :::::
8406 TATCTTGTAAACAGCAAGACGCTGCACATGAAGAT 8442

```

seq_name: /cgn2_6/plodata/2/lna/5A_COMB.seq:us-08-353-700-2

seq_documentation_block:

Sequence 2, Application US/08353700

Patent No. 5599919

GENERAL INFORMATION:

APPLICANT: YEN, TIMOTHY J.

APPLICANT: RATTNER, JEROME B.

TITLE OF INVENTION: NUCLEIC ACID ENCODING A

TITLE OF INVENTION: TRANSIENTLY-EXPRESSED KINETOCHORE PROTEIN,

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: DANN, DOREMAN, HERRELL AND SKILLMAN

STREET: 1601 MARKET STREET, SUITE 720

CITY: PHILADELPHIA

STATE: PA

COUNTRY: USA

ZIP: 19103-2307

COMPUTER READABLE FORM:


```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,700
FILING DATE: 09-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: REED, JANET E.
REGISTRATION NUMBER: 36,252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 563-4100
TELEFAX: (215) 563-4044
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 10136 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: HUMAN
US-08-353-700-2

```

```

alignment_scores:
  Quality: 158.50      Length: 2111
  Ratio: 0.171        Gaps: 93
  Percent Similarity: 44.008   Percent Identity: 17.859

```

```

alignment_block:
US-09-737-246-2 x US-08-353-700-2 ..

```

```

Align seg 1/1 to: US-08-353-700-2 from: 1 to: 10136

```

```

72 AspSerGlyProLeuArgAspLeuIleGluPheProAspIleG1 88
   ::::: ||| ::::: ||| ::::: |||
1950 GAGTCCAAAGCCTTCTGAGTCCTTACAGTTAAAAAGAAAGATATGA 1999
      88 uValValTyrSerProArgAspCysArgThrLeuValSerAlaValProG 105
      | ::::: ::::: ::::: ||| ||| :::::
2000 AGAATTG.....AAAAGAGAAACCTGTTCTGTGTTGGAAAA 2040
      105 LucIuSerGlu.....MetAspPro..... 111
      ::::: ||| ::::: ||| ::::: |||
2041 GTGAAAGCAAGAAACCTTTAACTCAGATGGAATCAGAAAGAAACCTTG 2090
      112 .....HisValArgAspCysIleArgSerTyr..... 120
2091 CAGAGTAAATTAATCACTTGGAACCTGTCTGAGACACAGCAATATAA 2140
      121 ..... 123
2141 AAGTCATGATATCAACAGAGAGAGTAAGACGCTGAGATGGACAGAGAAA 2190
      123 sPTPAlaIleValIleArgIleTyrHisIleValIleGlyThrGlyPheAsn 139
      ::::: ||| ::::: |||
2191 ACCTAAGTGTGAGATCAGAAACCTTCAC..... 2219
      140 ProAsnThrLeuAspIleValGlyLeu.....GluArgIleValGlyLe 153
      ||| ::::: ||| ::::: |||
2220 ...AACGCTTATGACATGAGTACAGTACAGTACAGTACAGTACAGTAC 2266
      153 uProIleValIleValIleValIleValIleValIleValIleValIleVal 170
      ::::: ||| ::::: |||
2267 TTATATGAGAGTACAGCAAGAAAGCTGAGTTCTCAGATCAGAAACATCAGA 2316
      170 LAspAspGln..... 173
      ::::: |||
2317 AGGAATATGAAATATGTGTTGAAGACTCTCAGCTTACTGGCAAGCTT 2366

```

```

174 AspAspLeuLysArgArgSerMetSerIleAspAspThrProArgIleGly 190
   ::::: ||| ::::: ||| ::::: |||
2367 GAGATCTTACAAACAGCAAGCTTCAGTACTGTCAATGAATATATGACAAA 2416
      190 rTPAlaCys.....SerIlePheAspL 198
      ::::: ||| ::::: |||
2417 AGACCGGTGTACCAAGACTTCGATCCGAAATGAGAGCCCTCAGAGATC 2466
      198 euLysAsnSerIleProAspAlaIleLeuProAsnIleuLysAspArgThr 214
      || ::::: |||
2467 TGTAAATTCAAAGATGCTTCTGTGTGACAAAT..... 2501
      215 ProAsnGluIleAspArgIleAsnAspAspGluArgIleLysSerAsnAr 231
      2502 .....GAAATCATCAGAGAAAGCTTTGGC 2527
      231 gHisLysGluLeuPheAlaLeuHisProSer..... 241
      ::::: ||| ::::: |||
2528 TTTTGATCACCACCCCTCCATCATTCCTTTGCCAATATATATGAGAG 2577
      242 .....ProAspGluGluGluProIleGluArgLeuSerVal 253
      ||| ::::: |||
2578 AACAGGAAGCATGCCCTTCAGAGAGAGTGAATGT...CGTTAGAGACA 2624
      254 ProAspIleProLysGluHisPheGlyGlnArgLeuValLysCysLe 270
      ::::: ||| ::::: |||
2625 GACCAAGTCCGAAAAAT.....TCTGCATCTCTCAAAATATGAGTTCA 2668
      270 uSerLeuLysPheGluIleGluIleGluProIlePheAlaSerIleAlaL 287
      ||| ::::: |||
2669 TTCACCTGAATTTTCAATTAAGTCTCAAAAACAGATCACTGACGCTGC 2718
      287 euTyrAspValLysGlu.....LysLysLysIleSerGluAsn 299
      ::::: ||| ::::: |||
2719 AAAAGCATGTGAAGAGTGTGTCAAATCAAAAGCAAAATAGAAAGAAAT 2768
      300 PheTyrPheAspLeuAsnSerGluGlnMetLysGlyLeuLeuArgProH 316
      ::::: |||
2769 CTC.....ATGAAGCAGAAACAGATG.....CA 2791
      316 sValProAlaAlaIleThrThrLeuAlaArgSerAlaIlePheSerI 333
      ::::: ||| ::::: |||
2792 TCAAGTTTGTGCTGAACAAAGTCAGCCGATTAAGTTACAGAGAG 2841
      333 IeThrTyrProSerGlnAspValPheLeuValIleLysLeuGlyVal 349
      ||| ::::: |||
2842 ACACCTTCTGCTCACACAGAAATGT..... 2864
      350 LeuGlnGlnGlyAspIleGlyGlyCysAlaGluProTyrMetIlePhe 366
      2865 .....GTGCTGA 2872
      366 sGluAlaAspAlaThrLysAsnLysGluLysLeu..... 380
      ::::: ||| ::::: |||
2873 AACCTTAAGTCCCTTGAGAAAGAAAGAGCGTCAACTTTAAATG 2922
      381 ..LysSerGlnAlaAspGlnPhe...CysGlnArgLeuGlyLysTyrArg 395
      || ::::: |||
2923 ATAAAGTAGGAACCTGAGCAGCAGAGATTCAGAAATTAATAAAG..... 2966
      396 MetProPheAlaTrpThrAlaIleHisLeu..... 405
      ::::: |||
2967 .....ACCAACCATCTACTGAAAGACTCTTAAGCA 2998
      406 MetAsnIleValSerSerAlaGlySerLeuGluArgAspSerThrGluV 422
      ::::: ||| ::::: |||
2999 GCTACACACTTTATCCGAACCTTAAGCTTGAGAGAAAGAAATGAGT 3048
      422 alGluIleSerThrGlyGluArgLysGlySerTrpSerGluArgAsn 438
      ||| ::::: |||
3049 CCATCATTTCTTAATAATTAAGGAAATGGAAGAGCTGACCAAGAGAT 3098

```

```

3735  ... AAATTTAAACCTCAGATGATCTGTGAAGTTAAAGAAATTTCT... 3776
730  eH1stHrGlnAspProTyrLeuAspIysPheHleAlaLeuValAsnAlaIu 747
3777  ..... CTGAGATGGTATATATGCGCAGTTGGTGCAT 3807
747  euAspGlnHisLeuPheProValATrgIleGlyAspMetArgIleMetGlu 763
3808  TAGAAGCTATGCTA.....AGAAATTAAGCAATTAAACTTCAGGA 3848
764  AsnAsn.....LeuGlnGluLeuLysSerSerIleSe 775
3849  AGTGAAGAGAGAGAGAGAGTCCCTGCAGCATGTAATTACAGACAAATAGAGG 3898
775  rAlaLeuAsnSerSerGlnLeuGlnProValValArgPheLeuHisLeu 792
3899  AGATCTTGAACCGAGCAATTTCCAGACACTGAG..... 3932
792  euLeuAspLysLeuIleLeuLeuValIleArgProProValIleAlaGly 808
3933  .....TCACA 3938
809  GlnIleValAsnLeuGlyGlnAlaSerPheGlnAlaMetAlaSerIleI 825
3939  GAATATTAGCGGCTTAAAGACCTGTGAATAAGATCGGAGAAACCTATAT 3988
825  eAsnArgLeuHisLysAsnLeuGlnGlyAsnHisArgGlnHisGlyArg 842
3989  TTGAGGCGCTCATGACTTGTCAACAAGTCAAAACACAAATGCACACCTTC 4038
842  snSerLeuLeuAlaSerTyrIleHisTyrValPheArgLeuProAsn... 857
4039  AGTGCCTCTGCACAACAACAATGAAGAGCTAAATAGACTAGAGAAATA 4088
858  .....ThrTyrProAsnSe 862
4089  TGTGAATATCTGCAGGCTGAAAAGTATGAACTGTAACTGAGCTGAATGA 4138
862  rSerSerProGlyProGlyGlyLeuGlnGlySerValHisTyrAlaThr 879
4139  TTCAGAGCTA.....GAAATGTATCAAG 4161
879  eTAlaArgSerAlaValArgProAlaSerLeuAsnLeuAsnArgSerArg 895
4162  CAACTGTGAATAATGGCAGAGAAGAGTAGGAAACTCTAAATGAAGTTAA 4211
896  SerLeuSerAsnSerAsn.....ProAspI 904
4212  ATTATTAAATGATCAGACGTGCTTCTTCATGCTGAGTGTAGTGAAGACAT 4261
904  eSerGlyThrProHisSerProAspAspGluValArgSerIleIleGly 921
4262  ACCAGAGAGTGTAATTTGGTGAACAACCAATTAACAGCACCCCTGTGCTT 4311
921  eTyrGlyLeuAspArgSerAsnSerTyrValAsn...ThrGlyGlyPro 936
4312  TGGCTTCATTTGGACGAGATATTTCTCTACGACACTTGTACATTGTGCAGC 4361
937  Lys.....AlaAlaProTyrGlySerAsnProSerPr 947
4362  AAAGAGATTCAATGCACATTTCGCAATTTGCAGAAGAAATTTCTTATCTTT 4411
947  oSerAlaGluSerThrGlnAlaMetAspArgSerLysAsnArgMetSer 964
4412  ACAAAATGAGACAAAAATTTTACATGATCAGACGACGTAGATGAGCTTGA 4461
964  eTyrHisThrLysThrSerSerPheLeuGlnThrLeuThrGlyArg..... 978
4462  AAATGTCAAGAGCTGCAGACCTTATTTGACTCTATTAAAGCCGAAATTTTG 4511
979  .....LeuProThrLys 982

```


Align seg 1/1 to: PCT-US95-16216-2 from: 1 to: 10136

2528 TTTTATATCAGACACCCGCGCATCATCATTCCTCTTGCAAAATATAATTGGAG 2577

242ProaspGluGluGluProIleGluArgLeuSerVal 253

2578 AACAAAGAAAGACATCCCTCAGAGAGAGATGATGT...CGTTTAGAACGA 2624

254 ProAspIleProIleGluGluHisPheGlyGlnArgLeuValValCysLe 270

2625 GACCAAAAGTCGCAAAAT.....TCTGCATCTCTPACAAATATAGATTGA 2668

2669 TTACTTGAATTTTCAATTAGAGCTCTCAAAAACAGATGAACCTACAGCTCG 2718

287 eutArgAspAllysglu.....LysAlaLysIleSerGluAsn 299

2719 AAAGAGAGTGTGAAGGTGGTGCGCAATCAAAAGAAAGAAATAGCAAAAT 2768

300 PheTyrPheAspLeuAsnSerGluGlnMetLeuTyrGlyLeuLeuArgProH 316

2769 CTC.....ATCAAAACAGAACAGATG.....CA 2793

316 sValProProAlaIleIleThrThleuAlaArgSerAlaIlePheSerI 333

2792 TCAAGATTGTTGGCGCGAAACAGTCAGCGCATGTGTAAGTTACAGAGAG 2841

333 IerhTyrProSerGlnAspValPheLeuValIleGlyLeuGluVal 349

2842 ACACCTCTCTGCACCGAGATGT..... 2864

350 LeuGlnGlnGlyAspIleGlyGluCysAlaGluProTyrMetIlePheLy 366

2865GTRCGTA 2872

366 sGluAlaAspAlaIleThrLysAsnLysGluLysLeuGluLysLeu..... 380

2873 AACCTTAAGTCGCCCTTGAAACAGAGAAAGAAAGCTGCAACCTTTAAATG 2922

381 ..LysSerGlnAlaAspGlnPhe...CysGlnArgLeuGlyLysTyrArg 395

2923 ATTAGGTAGAAACTGACGCGACAGATTCAGAAATTAATAAAG..... 2966

396 MetProPheAlaTyrPheAlaIleHisLeu..... 405

2967AGCAACCATCTACTGTGAAGCTCTTAAGCA 2998

406 MetAsnIleValSerSerAlaGlySerLeuGluArgAspSerThrGluY 422

2999 GCTACACACTTTATCCGAACCCCTAAGCTTGAGAGAAAGAAATGAGATT 3048

422 algluIleSerThrGlyLysArgLysGlySerTyrSerGluIlaArgAsn 438

3049 CCATCATTTCTTTAAATAATAAAGGAATTTGAAGACTGACCCAGAGAGAT 3098

439 SerSerIle.....ValGlyArgArgSerLeuGluArg 449

3099 GGGACCTTAAGGAATTAATGATCATCTTAATATCAAGAGAAAGATCACTT 3148

449 gThrThrSerGlyAspAspAlaCysAsnLeuThrSerPheArgProAlat 466

3149 AATCCAGAAAGATGAGAGATTGTCAAACCTATATATGATGAAGGAGAGAAA 3198

466 hIleLeuThr...ValThrAsnPhePheLysGlnGluGly..... 477

3199 GCATTTCAGAGTATCTGTGATCAGTACAAAGCAAGAAAACTTATTTTACTA 3248

478 AspArgLeuSerAsp.....GluAspLeu.....TyrLysPhe 488

3249 CAAGAGTGTGAAGAAACCGAAATGCAATATGAGATTTTACTCAAAATA 3298

488 eIeuAlaAspMetIaArgArgProSerSerValLeuArgArgLeuArgProI 505

3299 CAAGAGACAGACAGGAAAGAAATTTCTAAATTA..... 3329

505	lethlaaGlLeuLysIleAspIleSerProAlaProGluAsnProHis	521
3329	3329
522	tyrCysLeuThrProGluIleuLeuAlnValLysLeuTYrProAspSerat	538
3330	GAATGCTTCTAAATGAAATGCACCTAGTCTT.....TGTGAAATAG	3370
538	gvalArgProThrArgGluIleLeuGluPheProAlaArgAspValTYrv	555
3371	GAATAATGAGTGGAAACAGCTAAAGAGACATTTGCAAAAGAA.....	3413
555	alProAsnThrThrTYrArgAsnLeuLeuTYrIleTYrProGlnSerLeu	571
3414CACCAAGATTTCTTA.....ACAAATTA	3437
572	AsnPheAlaAsnArgGlnIleSerAlaArgAsnIleThrValLysVal..	587
3438	GCATTTGTCAGAAAGAAAT.....CAGATGTCAGTACAGTTGGA	3481
588GlnPheMetTYrGlyGluAspProSerAsnIleMetProV	601
3482	GACAGTCGACAGAGCTCTGAGATCGAAGATGACAGATAACCAAAACAT.	3530
601	allPheGlyIleSerSerCysSerGluPheSerLysGluAlaTYrTYr	617
3531TCTAAGACCGAGCGTGGTGTTAACCAAGAAATATGACT	3572
618	AlaValValTYrHisAsnArgSerPro.....AspPheHisG1	630
3573	TTAAAGAAAGCAACAAACCAATGCAAAAGAAAGTATGACTTATTACA	3622
630	uGluIleLysValLysLeuProAlaIleThrLeuThrAspHisHisLeuL	647
3623	AGGAAATGACACGCTGATGATAGGTATAGACTAAACAT.....	3662
647	eupPheThrPheTYrHisValSerCysGlnGlnLysAsnThrProLeu	663
3663GAATGTCAAAATCTGAGATCAGAACCAAT	3692
664	GluThrProValGlyTYrThrTrpIleProMetLeuGlnAsnGlyArgL	680
3693	AGGAACCTCTGT.....AAAGAAAGGA	3715
680	uLysThrGlnGlnPheCysLeuProValSerLeuGluLysProProGlnA	697
3716	GACTGAGCAAAATCAATGT.....	3734
697	LatYrSerValLeuSerProGluValProLeuProGlyMetLysTrpVal	713
3734	3734
714	AspAsnHisLysGlyValPheAsnValGluValAlaValAlaSerSerL	730
3735	...AATTTTAAACCTCAGATGAGATCTTGAGTTAAAGAAATTTCT....	3776
730	eHisThrGlnAspProTYrLeuAspLysPhePheAlaLeuValAsnAlaL	747
3777CTGAGATGTTTAAATCCGACAGTTGGTGCAT	3807
747	eAspTrpHisIleuPheProValArgGlyLeuAspMetArgLLeuMetGlu	763
3808	TAAACCTATGCTA.....ACAATTAAGCAATTAAACCTTCAGGAA	3848
764	AspAsn.....LeuGluAsnGluLeuLysSerSerLLeS	775
3849	AGTGACAAAGGAAGAGAGCTCGACAGCATGATTAACGACAAATTAGAG	3898
775	ralaLeuAsnSerSerGlnLeuGluProValValArgPheLeuHisLeuL	792
3899	AGATCTTTGAAACCGACGAATTTCCAGACATCGAG.....	3932

```

792 euleuaspIysleuIleuIleuValIleargProProValIlealagIy 808
3933 .....TCACAA 3938
809 GlnIleValasIleuGlyClnIlaSerPheGlnIleMetalaSerIleI 825
      :|||:| :|||:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
3939 GAATTAGTGGCGCTTAAGACACTGTGAATATAGTGGGAGAAAAGTATAT 3968
      :|||:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
825 easnargIleuIstIysAsnIleuGlyIAsnHisaspIlnIstcIylarG 842
      :|||:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
3989 TTCAGGCGCCTCATGAGTGTGTCAACAAGTCAAAAGACAAATGCACACCTTC 4038
      :|||:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
842 snSerIleuIleuIaSerTylIeIstIylValPheargLeuProasn... 857
      :|||:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
4039 AGTGCCTCTCGCAACACAAATGAACAAGCTGAATGACCTTAGAGAAATA 4088
      :|||:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
858 .....ThyTyrProAsnse 862
      :|||:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
4089 TGTGAATATCTCGAGCGTGAANAAGTATGTAACCTGTAACGTGAGTGAATGA 4138
      :|||:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
862 rSerSerProGlyProGlyGlyLeuGlyGlySerValHisTylIaThrM 879
      :|||:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
4139 TTCACAGCTCA.....GAATGTATCACAG 4161
879 etalIaargSerAlaValArgProIaSerLeuAsnIleuAsnargSerarg 895
      :|||:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
4162 CAACCTGAGAAATGGCAGACAGTAAAGGAAACCTACTAAATGAAGATTAAA 4211
896 SerLeuSerAsnSerasn.....ProaspII 904
      :|||:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
4212 AATATTAAATGATGACAGTGGCTTCTCCATGGTGAGTTAGTGAAGACAT 4261
904 eSerGlyThrProThrSerProAspGlyValaIargSerIleIleGlys 921
      :|||:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
4262 ACCAGAGAGGTGATTTTGGCAACAACAATGACAGCACACCTGTGTCTT 4311
921 etIysGlyLeuaspArgSerAsnSerTyrPvalasn...ThrGlyGlyPro 936
      :|||:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
4312 TGCGTCATTTGGCAGAGAGTAAATTCCTACAGACACTTGAATTTGCACAG 4361
937 Lys.....AlaIaIaProThrPglYSerAsnProSerPr 947
      :|||:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
4362 AAGGAAGTTCAATATGCATTTCCCAATTTGCAGAGAAATTTCTATCTTT 4411
947 oSerIaIaIuSerThrGlnIleMetAspArgSerGlyAsnArgMetSers 964
      :|||:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
4412 ACAAAAGTGAACACAAAATTTTCAATGATACAGACACTGTACATGACCTTA 4461
964 erHisThrGlnThrSerSerPheLeuGlnThrLeuThrGlyarg.... 978
      :|||:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
4462 AATATGCAGAGCTGCAGACCTATGTTGACTCATTTAAAGCCGAAAAATTTG 4511
979 .....LeuProThrIly 982
      :|||:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
4512 GTCCTTGCACGAATCTGAGAAATTTCAAGGTGACTTGGTGAAGAGAT 4561
982 sIysLeuPheHisGlnIu.....LeuIaIaLeuGlnTyrP 994
      :|||:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
4562 GCAGCTGGCGCTTGGAGAGAGGCGCTGTTCACCTCTGTCAATCCTTTGTG 4611
994 alValCysSerGlySerValArgGluSerAlaLeuGlnGlnAlaIaTrpHe 1010
      :|||:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
4612 TGCGTCAACAGCTCTACTCTT....AGCAGTTTTGGAGACGCTCCCTTT 4655
1011 PhePheGlnIleuMetValIysSerMetValHisIstIleuTyrPheasn. 1026
      :|||:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
4656 TACAGAGCTCTTTTGAACAACAGACAGAGATATGCTCTTTTGTAGATAATT 4705
1027 .AspIysIleuGlnAlaIaProArgIlySerIaIaPheProGlnIaIaIaPheMetA 1043
      :|||:|:| :|||:|:| :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
4706 AGAAGGGCGCTTTTCCGAACAACCGACGAGTGTAGATACGATATTTTGCCA 4755
1043 spaSpIleIaIaIaIeValIserThrIle...AlaSerAspIleValIser 1058

```



```

1526 valSerLysePheProGluLeuLeuPheGluGluGluThrGluGluCysAl 1542
      |||
6483 GTG.....AGGAAAGACCGAA..... 6500
1542 aaSPLeuCyLeuArgLeuLeuArgHisCysSerSerSerIleGlyThrI 1559
      |||
6501 .....CTCTTCAGACTTGTCTCTGATGTGAGTGAGC 6534
1559 le.....ArgSerHisProSerAlaSerLeuLeuLeuMetArg 1572
      |||
6535 TGTTAAAGACAAACTCATCTCCAGAAAGCTGCAGAGTTGGAAAG 6584
1573 GluAsn.....PheGluIleGlyAsnAsnPheAl 1582
      |||
6585 GACTCACAGGCACTGCTTTGACAAATGTGAGCTGGAACCAATTC 6634
1582 aaArgVal..... 1584
      |||
6635 ACAACTGAATTAAGAGAAAGAAATTCCTGTCAAGAAATCGAAAGCTGC 6684
1585 .....LysMetGlnValProMetSer 1591
      |||
6685 AGCCCACTGACTGATGATGATTAAGAAAGCTGATGCTCCAGAGCC 6734
1592 LeuSerSer.....LeuValG1 1597
      |||
6735 TTGGAGCGCCGCACTGCTGAGAAAGTGCAGTTCCGATTTAGGCTGAGC 6784
1597 yThrSerGlnAsnPheAsnGluGluPheLeuArgArgSerLeuLysThrI 1614
      |||
6785 AACACAGAGAGAGAGTGCATCAG.....CTGAGAAAGAGCATCGAGAAAC 6828
1614 leLeuThrTyrrAlaGluGluAsp..... 1621
      |||
6829 TGAGACTGCGATGTGAGCCGATGAAAGACACAGCTGCATCCGACAG 6878
1622 ..LeuGluLeuArgGlu.....ThrThrPheProAspGlnValG1 1634
      |||
6879 AAACGTAAAGAAAGCGGAGCGGAGATGATTCACTTAAGATTAAGTGA 6928
1634 naSPLeuValPheAsnLeuHisMet.....IleL 1644
      |||
6929 GAACCTTGAAGAGGATTCACATGTCACAGAAACCAAGAGCTAGTGA 6978
1644 euserAspThrValLysMetLysGluHisGluAspProGluMetLeu 1660
      |||
6979 TCTTGATGCCGAAATTCCAAGCGAGAGCTGACACTTAATAAACACAA 7028
1661 IleAspLeuMetTyrrArgIleAlaLysGlyTyrrGln..... 1672
      |||
7029 ATAGAAGAGATGCGCAGAGCCTGAAGATTTTGAATTAGACCTTGTCAC 7078
1673 .....ThrSerProGluArgLeuThr...TrrLeuGlnAsnMetAlaG 1686
      |||
7079 GTTAAGCTGTGAAAAAAGAAATCTCACAACCAATACAGAAAAAACAAG 7128
1686 lYlshHisSerGluArgSerAsnHisAlaGluAlaGluCysLeuVal 1702
      |||
7129 GTACAGTTGCACAACTAGCAAG..... 7151
1703 HisSerAlaAlaLeuValAlaGluTyrrLeuSerMetLeuGluAspArg 1719
      |||
7152 .....TTACTCTCTTCATTTAAAGTCTGTAGAGAAAGAGA 7189
1719 styrrLeuProValGlyCysValThrPheGlnAsnIleSerSerAsnVal 1736
      |||
7190 GCAAGCAGAGATA.....CAGATCA 7209
1736 euGluGluSerAlaValSerAspAspValValSerProAspGluGluCyl 1752
      |||
7210 AAGAAAGATCTTAAACTGCATGCGAGATGCTT..... 7241

```

```

1753 IleCysSerGlyLysTyrrPheThrGluSerGlyLeuValGlyLeuG1 1769
      |||
7242 .....CAGATCAGTTAAAGGAGCTTAATGA 7267
1769 uGlnAlaAlaIleAserPheSerMetAlaGlyMetTyrrGluAlaValAsnG 1786
      |||
7268 GGCAGTAGCACCC.....TTGTGCGTGACCAAGAAATTTATGAAG 7308
1786 luValTyrrLysValLeu...IleProIleHisGluAlaIleAsnArgAspAla 1801
      |||
7309 CCACAGACAGAGCTTAGACCCACCACCAATAGAGAGAGACATCAGCTGAGA 7358
1802 LysLysLeuSerThrIleHisGlyLysLeuGln 1812
      |||
7359 AATAGCATTGAAAGAGCTGAGAGCCGCCCTAGAA 7391
seq_name: /cgn2_6/ptodata/2/lna/5A.COMB.seq:us-08-471-119A-1
seq_documentation_block:
; Sequence 1, Application US/08471119A
; Patent No. 5827706
; GENERAL INFORMATION:
; APPLICANT: Leitner, Ernst
; APPLICANT: Schneider, Elisabeth
; APPLICANT: Schoergerdorfer, Kurt
; APPLICANT: Weber, Gerhard
; TITLE OF INVENTION: Cyclosporin Synthetase
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5827706artis Corporation
; STREET: 59 Route 10
; CITY: East Hanover
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07936
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,119A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kassenoff, Melvyn
; REGISTRATION NUMBER: 26,389
; REFERENCE/DOCKET NUMBER: 100-8029/CONT/CONT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 503 8474
; TELEFAX: 201 503 8807
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46899 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Tolypocladium niveum
; STRAIN: ATCC 34921
US-08-471-119A-1

```

```

alignment_scores:
  Quality: 158.50      Length: 1789
  Ratio: 0.209         Gaps: 95
  Percent Similarity: 42.314      Percent Identity: 19.955
alignment_block:
US-09-737-246-2 x US-08-471-119A-1  ..

```



```

13740 CAGCTGCAATCTGGGTGAGCAACTGTAGATAGTCCCCGGCCGAGCT 13789
13789 hegluargmetasnserleuthrphelylsyserserlyaspmetargala 1394
13790 ACTCAGCATC.....TCGCCGGCCTTCGATCC 13818
1395 lysleugluajalaileuglyserileglvalaargln..... 1408
13819 TCTCTGTCACAG.....CTGGGAGATGTGCGACGATCGACGAGGACA 13862
1409 .....glumetvalaargargserargglygluargserpros 1423
13863 CTATACAAAGACTTGGAGAAATCTGCGGGTCATCGCCTTACCTCCTT 13912
1423 er.glyseralaphneglysergluasn.....leuargtr 1435
13913 CGTGTACTGCTTGCGGCCCTACCGCGCAGCCCATTCGCTCACAGTT 13962
1435 parglyaspmetrhrthrpararglnasnthrghuylusleuasplys. 1451
13963 CCGA.....AGACGCACTATAGGAGACACCAATCGCC 13994
1452 Serargalaagluilegluhisgluialeuileasprglyasnleuala1th 1468
13995 AATCTTACCGCACTTGACATGACGATATATCGCCTTCTTCGCAATAC 14044
1468 rglualaasnleuileileuaspthrleugluilevala1glnthr 1485
14045 GCATGTCTACGCATT.....KCCG 14064
1485 alservalthrghuserlysgluserileleuglyvalleuylsval 1501
14065 TCACGAGAGAGAGACCTTGATGCA..... 14090
1502 leuenuhissermetala1cysasnleu1a1val1tyrleu1gln.hisc 1518
14091 CTAGTGCAGCAGGTCCGGTCAACGCGCACAGCTCATTCGCCCATCAGGA 14140
1518 yspheala1thghlpargala1leuvalserlyspheproglu1leu1phe 1534
14141 CGTCCCGTTCGAGAGATCGTCTACTCTTTTGGCGTTCTCGAGATG 14190
1535 gluglu1guthrghlucysala1asp1leu1cysleu1argleu1leu1argh1 1551
14191 CATCCGGAACCCACTTGTGCAGCTCA...TGTTGGCGTGCATTCGAG 14237
1551 s..... 1551
14238 AAGAACTCGGTGAGCTGAGCTGGAACCTGGAACGCTCACAGCGAGTTGTCC 14287
1552 .....CysserSer 1555
14288 CACGGAGATCACAGCCGCTTGATTTGGAATTCACCTGTTCCAGCAG 14337
1556 ilegl1ythr1leu1argser.hisproserala1serleu1tyrleu1met 1572
14338 ATGACAAAGCTTGAAGGCTCCATCT.....CTATTCA 14369
1572 rgglu1asn.phegluilegl1yasn1asn1phe1a1argval1ysmetgl1na 1588
14370 ACCGATCTCTTCGAA..... 14384
1588 l1prometserleu1serleu1val1glythrserglu1asn1pheasn1glu 1605
14385 .GCATCTCGGTCCAAAGCTTCTTTCAGTATCCAG..... 14421
1605 l1uphe1leu1arg1argserleu1ys.....Thr1leu1thr1tyr 1617
14422 AATTTCTGCGCGGGGTTTGAACGCTCGGAGCTGCCCATCAGCACCTTA 14471
1618 Ala1glu1uasp.....leu1glu1le 1624

```

```

14472 CCACCTTCAGAGTGAATTCGTCGACCTACAAAGACAGAGCCCTGTGGATGT 14521
1624 uarglu1uthr1thr1phePro.....Asp1glu1a1gln1asp1leu1a1phe 1639
14522 CCAGAGACGCAATATCCCTCGATTCCTCTGTGTGATGTGTCCAG 14571
1639 snleu1hismet1leu1serasp1thr1val..... 1648
14572 AGCAGGTCTGCATCAACCCGATTCATTCACCTGATATCATGCTCGAG 14621
1649 lysmet1lysglu1hisglu1uasp1proglu1metleu1leu1asp1leu1met 1665
14622 AAGCTCAGCTACGCCCAAGCTCGACAGGAA.....TGTGA 14656
1665 rarg1leu1a1arg1tyr1gln1thrserproglu1argleu1thr1tyrleu 1682
14657 CAGGTCGCTCGC.....TGGCTC. 14675
1682 l1asn1metala1gly1ys1hisserglu1argserasn1h1a1glu1a1a1a 1698
14676 .....CGTCAACGTTCTTTCAGCTCCGACACGTTAATCGCG 14711
1699 glncysleu1val1hisserala1ala1leu1a1a1glu1tyrleu1sermetle 1715
14712 GTGTTGGCGCCACGGTCTTGCGAGACGATATCGCGTTCCTCGGAATCT 14761
1715 uglu.....Asparglysty1leu1proval1glycysval1thr1phe1asn1 1731
14762 CAAGGCAAACCTTCGCTACCTCTG..... 14789
1731 le1ser1asn1val1leu1glu1leu1serala1val1serasp1a1val1ser 1747
14790 .....GATGTCAAGGCTCTGCTGCCGCAATGATGCATTTATGCG 14831
1748 .....Proasp1glu1glu1tyr1leu1cys1serglu1tyr1thr1phe1glu1se 1762
14832 TCGCTACCCCGGAACAAGCTCATCTGCTGGCGCAACGTTACGCGGCC 14881
1762 rglyleu1val1glyleu1leu1glu1a1a1a1a1ser1phe1ser1metala1g 1779
14882 C.....AAGCTTCAGGAAGCGGCCATCGATTC..... 14909
1779 lymetyrglu1a1a1val1asn1glu1a1tyr1lysval1leu1leu1leu1 1795
14910 .....GTGCCATC... 14918
1796 glula1asn1arg1asp1ala1lys1leu1ser1thr1leu1hisglu1lysleu1 1812
14919 .....CGTATACCTTCACCTACACTCACT...GACGGCACACTTCA 14956
1812 nglu..Alaphe1ser1ys1leu1a1hisglu1ser1thr1gly1tyr1glu1rme 1828
14957 AGATGGCCCTAC.....CATCGAGCAACCTCTGCGCAAGCC 14994
1828 t1phe1gly1thr1tyr1phe1arg1val1gly1phe1tyr1glu1tyr1lys1phe1 1845
14995 TAGGTATCGGCATGTCACGTCGTGTTCTACCGGACGAGCAAGGGTGTT 15044
1845 eu.....Asp1glu1gln.glu1phe1a 1851
15045 ATGGTCCAGCAGCGCAACATCGTCCGCTGTGTAAGAAACAGTAACTCGT 15094
1851 l1tyr1lys1glu1pro1ala1leu1thr1lysleu1a1a1glu1leu1ser1his 1865
15095 CGCTAAGCAGCCCGCAGCAGAGCTCGCATAGCACATATATTCAT 15137
seq_name: /cgn2_6/p1odata/2/1na/5B_COMB.seq:US-08-457-273b-7
seq_documentation_block:
; Sequence 7, Application US/08457273B
; Patent No. 584995
; GENERAL INFORMATION:
; APPLICANT: Hayden, Michael

```


2424 2424
 372 sAsnLysGluLysLeuGluLysLeuLysSerGlnAlaAspGlnPheCysG 389
 2424 2424
 389 LnArgLeuGlyLysTyrArgMetProPheAlaThrPheAlaIleHisLeu 405
 2435TGG.....CTG 2430
 406 MetAsnIleValSerSerAlaGlySerLeuGlnArgAspSerThrGluVal 422
 2431 GGCAACATCAGAACCCCTGACAGGAAT 2457
 422 LgluIleSerThrGlyLysArgLysGlySerTrpSerGluArgAsn 439
 2458ACAT 2461
 439 eSerIleValGlyArgArgSerLeuGluArgThrThrSerGlyAspAsp 455
 2462 TTTCCTGCTGACTGCATTCCTTACTGACAGAAACCTGAGAGATGAA 2511
 456 Ala.....CysAsnLeu.....ThrSerPheArgProAlaThrLe 467
 2512 TCTTGTGTTACTTGCAGAGTTGGCTTGTAAGCTGTGAGGCACCTGTGCT 2561
 467 uThrValThrAsnPhenPheLysGlnGluGlyAspArgLeuSerAspGlu 484
 2562 GAGCTTTGGACGAGCAGCTACAGTGTGCTGG 2593
 484 sPheuTyrLysPheLeuAlaAspMetArgArgProSerSerValLeuArg 500
 2594 GATTA..... 2598
 501 ArgLeuArgProIleThrAlaGlnLeuLysIleAspIleSerProAlaPr 517
 2599CACTGCTTATGATATGATGCTGCCTGCA 2627
 517 OgluAsnProHisTyrCysLeuThrProGluLeuGlnValLysLeuT 534
 2628 GAACAGCTCCTACTGCTGTGAGAGCCGAAGCTGCTGACACTGTGGAG 2677
 534 yrProAspSerArgValArgProThrArgLuuIleLeuGlnPheProAla 550
 2678 AGATTGACTTTCAGGCTC.....GTGAGTTTTCGAGGCAAAAGCA 2718
 551 ArgAspValTyrValProAsnThrThrTyrArgAsnLeuLeuTyrIleT 567
 2719 GAAAGTTTACACCGAGGCTCATATTATACAGGTTTCTAAAGCTACA 2768
 567 rProGlnSerLeuAsnPhenAlaAsnArgGlnGlySerAlaArgAsnIleT 584
 2769 AGACGAGTACTCATA 2785
 584 hValLysValGlnPheMetTyrGlyLysPro..... 595
 2786 AGTGTGATATTATTGCTTGAGATGAAGCCAGGTTTCGACATGT 2835
 596SerAsnAlaMetProValIlePheGlyLysSerSe 607
 2836 GCTGCAACATCTTTACAAAGGCTTCCCAAGCTGTTTACAA..... 2880
 607 rCysSerGlnPheSerLysGluAlaTyrThrAlaValAlaTyrHisAsn 624
 2881 .TGTACCAAGACAGCTGATCCAGTTGTGGCTGTACGAGGATCGA 2929
 624 rGSerProAspPheHisGluGluIleLysValLysLeuProAlaThrLeu 640
 2930 GCAGTGTCTACCTGAGACCTCTCATGATGAG 2961
 641 ThrAspHisHisLeuLeuPheThrPheTyrHisValSerCysGlnG 657
 2962 ACCCAGCACCATCAC..... 2977

657 LnLysGlnAsnThrProLeuGluThrProValGlyTyrThrTrpIlePro 673
 2978ACTTTTCTGTGACACCA..... 2995
 674 MetLeuGlnAsnGlyArgLeuLysThrGlyGlnPheCysLeuProValSe 690
 2995 2995
 690 rLeuGluLysProProGlnAlaTyrSerValLeuSerProGluValProL 707
 2996TCACAGAA..... 3004
 707 eupProGlyMetLysTrpValAspAsnHisLysGlyValPheAsnValGlu 723
 3005TCTATAGAG 3013
 724 ValValAlaValSerSerIleHisThrGlnAspProTyrLeuAspLysPh 740
 3014 GCTATACCTTACTGCCAAGAAATAACAGATGTCAACCATGG..... 3052
 740 ePheAlaLeuValAsnAlaLeuAspGluHisLeuPheProValArgIleG 757
 3053AAACAATCTCTCAAGAGTTGTCGG..... 3079
 757 LysAspMetArgIleMetGluAsnAsnLeuGluAsn..... 768
 3080CAGTTTCTCATGACTCATTCAGCAACAACAGGCGACTCAC 3121
 769GluLeuLysSerSerIleSerAlaLeuAsnSerSe 780
 3122 ATTTGATGCTGTGAAGCCTTGCTCTCTCAGACGCTTTCAGCTT 3171
 780 rGlnLeuGluProValValArgPheLeuHisLeuLeuAspLysLeuI 797
 3172 GCAGTTGGAG..... 3181
 797 IeLeuLeuValIleArgProProValIleAlaGlyGlnIleValAsnLeu 813
 3182TTTAGG.....ANGCACAG 3196
 814 GlyGlnAlaSerPheGluAlaMetAlaSerIleIleAsnArgLeuHisL 830
 3197 TGGAGTGTCCCCCAGCTAGTCCCTGTGATGAGTCCAGAGAGCTGCAC.. 3244
 830 sAsnLeuGluGlyAsnHisAspGlnHisGlyArgAsnSerLeuLeuAlS 847
 3245 .TGTGGGATGGCTCCATGATTTCTCAC.....CTTG 3275
 847 eTyrTleHisTyrValPheArgLeuProAsnThrTyrProAsnSerSe 863
 3276 CTTTCATCAGCTTGTT.....CCCACTGATCT 3304
 864 SerProGlyProGlyLysLeuGlyGlySerValHisTyrAlaThrMetAl 880
 3305 CTCAGCCCATCAGAGATGCCCTTGATTTGGCTGGAAACTTGACAGCA 3354
 880 aArgSerAlaValArgProAlaSerIleAsnLeuAsnArgSerArgSerL 897
 3355 GTGCCCCCAAGTCTCTGAGAGTTTCATGAGCTCTGAGAGAAAGCCAAC 3404
 897 euserAsn.....SerAsnProAspIleSerGly..... 906
 3405 TCAGCAGCCACCAAGACAGAGAAATCTGCCCGCTGTGGGGATGGAGC 3454
 907ThrProThrSerProAspAs 913
 3455 TCTAGTCCCTTGTGTGAGAGCTTTTCTCCCACTGCTGAGAGGTATCA 3504
 913 rGluValArgSerIleIleGlySerLysGlyLeuAspArgSerAsnSerT 930
 3505 ATATCTGTGCTCATGTCTTGA.....CGATGTGACTCCT 3539

4998 CCGTGGACATGCTTTTGGCGAGTATGTTCACTCCACACACATGGCA 5047
1487 ValThrIuSerLysGluSerIleLeuGluGlyValLeuLysValLeu 1503
:::|||||:::|||||:::|||||
5048 TCTGTAACACATGTGCAGACGTGTGATATCTGGAATCTCTCCCACTTCTG 5095
1503 uHisSerMetAlaCysAsnGlnSerAlaValTyrLeuGlnHisCysPheA 1520
:::|||||
5096 ...AGGGTCTCATTTCCAGTCA..... 5116
1520 IaThrGlnArgAlaLeuValSerLysPheProGluLeuLeuPheGluGlu 1536
|||||:::|||||:::|||||
5117 ...ACGGAGACATTTGTTGTCTGATTCAGAGCTCTCTCTCTCCCA 5164
1537 GluThrGlnGlnCysAlaAspLeuCysLeuArgLeu..... 1548
:::|||||:::|||||
5165 CACTGTCTCTCTCTGTTCCAGTGATTAACAGGTTAAGGGGTGGAGGCGG 5211
1549LeuArgHisCysSerSerSerIleGlyTyrIleArgSerH 1562
|||:::|||||:::|||||
5212 TAATGTAACACTAGAGAAATGCAGCGAA.....GGGAAACAAAGAGTT 5255
1562 IsProSerAlaSerLeuTyrLeuLeuMetArgGlnAsnPheGluIleGly 1578
|||||:::|||||
5256 TGGCAGAA..... 5263
1579 AsnAsnPheAlaArgValLysMetGlnValProMetSerLeuSerSerIle 1595
:::|||||:::|||||:::|||||
5264 GATACATTTCTCAGGTTCTTTTCTTACAGCTGGTGGATTTCTCTAGAAAG 5313
1595 uValGlyTyrSerGln.....AsnPheAsnGluGlu.....P 1606
:::|||||:::|||||
5314 CATGCTTACAAACAGCTCAAAAGTGACATGAGTGAACACACACATACGT 5363
1606 heLeuArgArgSerLeuLysThrIleLeuThr..... 1616
|||:::|||||:::|||||
5364 TCTACTGCCAAGAGCTAGGCACACTGCTCATGTCTGATCCACATATTC 5413
1616 1616
5414 AAATGGAATGTTCCGAGAAATCACAACAGCTCCACATGACCTTCCAC 5463
1617TyrAlaGlnGluAspLeuGluLeu 1625
5464 CAGTATGCTGTGTAAGGACACTTCTATCTGTAGAGAGCTGATATGCAC 5513
1625 TGGTThrThrPheProArgGlnValGlnAspLeuValPheAsnLeuHis 1641
|||:::|||||
5514 GGGTCCGATCCATGGTGGCCGACGCCACGCCCTGGTA..... 5551
1642 MetIleLeuSerAspThrValLysMetLysGlnHisGlnGluAspProGlu 1658
:::|||||:::|||||
5552 CTGGCTGTGTCAAGATCTCTACTCTCATCAACACACTGCAC..... 5593
1658 uMetLeuLeuAspLeuMetLysTyrGlnIleAlaLysGlyTyrGlnThrSerP 1675
|||||
5594TACCGTGTGGGCGACAGAGTGCAGCAGACAC 5624
1675 TGGTLeuArg.....LeuThrTyrLeuGlnAsnMet.....Ala 1685
|||||:::|||||
5625 CCAAGAGACACAGTCTGTCTGCAGAGTCACTTAACCCCAACATCT 5674
1686 GlyLysHisSerGluArgSerAsnHisAlaGluAlaLagLys..... 1700
5675 GGGCAGAGAGAGATTTGGCTCGCAGCTCAGTGGGAATGTCCATAAG 5724
1701LeuValHisSerAlaAlaLeuVal.....AlaGluTyrLeuSerH 1714
:::|||||:::|||||
5725 AGAATAATAGCGGAGAGGGCCCTTATTTCTCTGTGATTAATGTCTGTC 5774
1714 et...LeuGluAspArgLysTyrLeu..... 1721
:::|||||:::|||||
5775 AGAATCTCATGACTCAGAACACTTAACATGCTCATTTGTGAATCAGATT 5824

1722ProValGlyCysValThrPh 1728
5825 CAAGATCTGATCAGCTTGTCTCATGAGCTCCAGTA..... 5860
1728 eGlnAsnIleSerSerAsnValLeuGluGlnSerAlaValSer..... 1742
|||||:::|||||
5861 .CAAGACTTTATGATGTCATTCATTCGTAATCTCGCAGTACGTCTTT 5909
1743AspAspValValSerProArgGlu 1750
5910 TTATCCAGGCAATTCAGTCTGCTGTAATAATCTTCAACGCCAACACT 5959
1751GluGlyIleCysSerGlyLysTyrPh 1759
5960 CTGAAGAAACACTTTCAGTCTTGAAGGCATC.....CATCT 5997
1759 eThrGluSerGlyLeuValGlyLeuLeuGlu.....GlnAlaA 1772
5998 CAGCCAGTCTGGCCCTGTCTCACACTTATATGTGCAGAGCTCTGGGCA 6047
1772 IaAlaSerPheSerMetAlaGlyMetTyrGlu.....AlaValAsnGlu 1786
6048 CCTCTCCGCTGGCCTGTGCTGCATGTGCACACCCCTGGCTGTGCGCG 6097
1787 ValTyrLysValLeu.....IlePr 1793
6098 GTGAAGATGCTTTTGGCTGCAAAATTTACAGACAGCATGCCCAAGTTGCC 6147
1793 OIleHisGlnAlaAsnArgAspAlaLysLysLeuSerThrIleHisGlyL 1810
6148 AGAGAGGlnAlaAsnArgAspAlaLysLysLeuSerThrIleHisGlyL 1816
1810 yLysGlnGlnAlaPheSerLysIleValHisGlnSerThrGlyTyrGlu 1826
6177 ACCTCCAAACAGTGGGCTTGCACAAAGACCCAA..... 6211
1827 ArgMetPheGlyThrTyrPheArgValGlyPheTyrGlyThrLysPheGlu 1843
6212 AGGCTCTATTCAGTCTGCAGACA.....TTCCG 6240
1843 yAspLeuAspLysGlnGluPheValTyrLysGluProAlaIleThrLysL 1860
6241 ACCTCTACTGTGCAGACTCATACCCCTTGCACCCCACT..... 6286
1860 euAlaGluIleSerHisArgLeuGlnGluPheTyrGlyGluArgPheGly 1876
6287TCCACCCCACTGGGTGG.....GATGGG 6310
1877 GluAspValValGluValIleLysAspSerAsnProValAspLysCysLys 1893
6311 CACACATCTCGAAGAACAGT..... 6331
1893 sLeuAspProAsnLysAla...TyrIleGlnIleThrTyrValGluPro 1909
6332AGTCCAGACAAAGACTGTACTCTCAGCTTGTCAATCCAGTGT 6377
1909 yF.....PheAspThrTyrGluMetLysAspArg 1918
6378 GGACAGATCAAGATTTCTCACTGCTGGAAGGTGCAGAGCTGGTCAACCGT 6427
1919 IleThrTyrPheAspLysAsn.....TyrAsnIle 1928
6428 ATCCCTGCTCAGAGATGATGACTTCAATGATGAGCTCCGAGTTCAACCT 6477
6478 AAGCTTTTGGCTCCCTGTTAAGCTTGGCATGAGGAGATTTGCTAATG 6527
1945 yLysIleuHisGlnGlnPheLysArgLysThrIleLeuThrThrSerHis 1961
6528 GCCAAAGAGTCCCTCTTGAAGCA..... 6553

[illegible]


```

375 LuLysLeuGluLysLeu.....LysSerGlnAlaAspGlnPhe 387
      :::::|||||:::
1430 GCGCGCTGTGAGCTGCTGCGCTATCATCTCCAGCTCGAGATGAG 1479
      :::::|||||:::
388 CysGlnArgLeuGlyLysTyrArgMetProPheAlaTrpThrAlaIleH 404
      :::::|||||:::
1480 ATGACATCTCTGTCTAGAGTAT..... 1500
404 sleuMetAsnIleValSerSerAlaGlySerLeuGlnArgAspSerThrG 421
      :::::|||||:::
1501 .....GTTCTCGCATGAGAGACACAGAGT 1528
421 luValGluIleSerThrGlyLysArgGlySerTrpSerGlnArgArg 437
      :::::|||||:::
1529 CCATCTATTACATCATCTGTGTGAGACAAA.....GAGAGGTGGCC 1569
438 AsnSerSerIleValGlyArg.....ArgSerLeuGln..... 448
      :::::|||||:::
1570 AACTCAGCTTTGTGTGAGCGAGCGGAAACGGGCTTCGAGGTGTATTA 1619
449 .ArgThrThrSerGlyAspAspAlaCys.....AsnLeuThrSerPhe 463
      :::::|||||:::
1620 TATGACCGAGCCCATGTCAGCATCTGTGTGAGCAGCTCAAGAAATTG 1669
463 rgrProAlaThrLeu...ThrValThrAsnPhelLysGlnGluLys 478
      :::::|||||:::
1670 ATGGAGAGAGCCTGGTCTCAGTTACC.....AAGAGGGCTCG 1707
479 ArgLeu...SerAspGluAspLeuTyrIrisPheLeuAlaAsp.MetArg 494
      :::::|||||:::
1708 GAGCTGCGCTGAGATGAGAGAGAGAAGAAGATGAGAAAGAGCAAGC 1757
494 rgr.....ProSerSerValLeuArgArg 501
      :::::|||||:::
1758 AAAGTTTGAGAACCTTCAGACTCATGAAGAATCTTACATAGAGAG 1807
502 LeuArg.....ProIleThrAlaGlnLeuLysIle...AspI1 513
      :::::|||||:::
1808 TTGAGAGAGTGACAAATCTCCATAGACTGTGTCTTCACTGCGCAT 1857
513 eSerProAlaPro.....GluA 519
      :::::|||||:::
1858 GTGACGAGCAGCTACGCGTGACAGCAATATGAGCGGATCATGAAGC 1907
519 snProHisTyrCysLeuThrPro..... 526
      :::::|||||:::
1908 CCAGGCACTTCGGAGACACTCCACCATATGATGCGCAAAAAGC 1957
527 .....GluLeuLeuGlnValLysLeuTyrPro.....AspSerAr 538
      :::::|||||:::
1958 ACCTGAGATCAACCTTGACACACCCCATTTGTGAGACGCTGGGAGAA 2007
538 gValArgProThrArgGluIle..... 545
      :::::|||||:::
2008 GCTGAGGCGGACAGATATGATTAAGACCTGCTGTCTCTCT 2057
546 .....LeuGluPheProAlaArgAspValLys 554
      :::::|||||:::
2058 GTTTGAACCGCCCTGCTATCTTCTGCTTTCCTTGAGG..... 2098
555 ValProAsnThrThr 559
      :::::|||||:::
2099 ATCCCGACGCCACT 2113
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq.US-08-629-001A-2
seq_documentation_block:
; Sequence 2, Application US/08629001A
; Patent No. 5858661
; GENERAL INFORMATION:
; APPLICANT: Shion, Yosef
; TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE AND ITS
; ; ; GENOMIC ORGANIZATION

```

```

; ; ; NUMBER OF SEQUENCES: 139
; ; ; CORRESPONDENCE ADDRESS:
; ; ; ADDRESSEE: Kohn & Associates
; ; ; STREET: 30500 No. 5858661thwestern Hwy.
; ; ; CITY: Farmington Hills
; ; ; STATE: Michigan
; ; ; COUNTRY: US
; ; ; ZIP: 48334
; ; ; COMPUTER READABLE FORM:
; ; ; MEDIUM TYPE: Floppy disk
; ; ; COMPUTER: IBM PC compatible
; ; ; OPERATING SYSTEM: PC-DOS/MS-DOS
; ; ; SOFTWARE: PatentIn Release #1.0, Version #1.30
; ; ; CURRENT APPLICATION DATA:
; ; ; APPLICATION NUMBER: US/08/629,001A
; ; ; FILING DATE:
; ; ; CLASSIFICATION: 435
; ; ; ATTORNEY/AGENT INFORMATION:
; ; ; NAME: Kohn, Kenneth I.
; ; ; REGISTRATION NUMBER: 30,955
; ; ; REFERENCE/DOCKET NUMBER: 2290,00032
; ; ; TELECOMMUNICATION INFORMATION:
; ; ; TELEPHONE: (810) 539-5050
; ; ; TELEFAX: (810) 539-5055
; ; ; INFORMATION FOR SEQ ID NO: 2:
; ; ; SEQUENCE CHARACTERISTICS:
; ; ; LENGTH: 9171 base pairs
; ; ; TYPE: nucleic acid
; ; ; STRANDEDNESS: single
; ; ; TOPOLOGY: linear
; ; ; MOLECULE TYPE: cDNA
; ; ; ORIGINAL SOURCE:
; ; ; ORGANISM: Homo sapiens
; ; ; POSITION IN GENOME:
; ; ; CHROMOSOME/SEGMENT: 11q22-23
; ; ; US-08-629-001A-2

alignment_scores:
Quality: 150.50 Length: 2063
Ratio: 0.175 Gaps: 99
Percent Similarity: 41.638 Percent Identity: 17.402

alignment_block:
US-09-737-246-2 x US-08-629-001A-2 ..

Align seg 1/1 to: US-08-629-001A-2 from: 1 to: 9171

267 ValLysCysLeuSerLeuLysPheGluIleGluLeuProIlePheAl 283
      ||| |||||:::|||||:::
1615 GTATGCTGTGTGACTTTGGCAGCTGACGACGAGTATGCTTCAGGAACG 1664
283 aSerLeuAlaLeuTyrAspValLysGlyLysLysIleSerGluAsn 300
      :::::|||||:::
1665 AAAATGGGATA.....GAGCAAAATATGTGTGAAGTAATA 1702
300 heTyrPheAspLeuAsnSerGlnGlnMetLysGlyLeu..... 313
      :::::|||||:::
1703 GAACCTTTCTTTAAAGAAATCAATATGAATGAGCTCTATTTCTATCAG 1752
314 .....ArgProHisValProProAlaAlaIleTh 323
      :::::|||||:::
1753 TTAGAGGTGCTAGAAATATGACACAGAAAGTGCCTCA..... 1791
323 rThrLeuAlaArgSerAlaIlePheSerIleThrTyrProSerGlnAsp 340
      :::::|||||:::
1792 .....ATTCTCACAGTAAATTTCTCTCAT..... 1815
340 alPheLeuValIleLysLeuGluLysValLeuGlnGlnGlyAspIleGly 356
      :::::|||||:::
1816 .....CTTGACTGTGAGAAATCTTGTGAGACTTCACACTATGAAA 1854
357 GlucyAlaGluProTyrMetIlePheLysGluAlaAspAlaThrLysAs 373

```

2470	TCCTTATGCAAAAAGCCATTGACCGTGAGAGTAGAATCAATGGAAGA	25119
665	rProValIGlyTYrThrDLeProMetLeuInsnGlyArgLeuLys.	681
2520	TGATACT.....AATGGAATCTGAATGG	2542
682ThrGlyInPheCysLeuProValSerLeuGlnLysProGln	696
2543	AGGTGAGAGATCAGTCAATCCAAAGATCTATTAAAGATTACCTGATAGT	25922
697	AlaTySerValLeuSerProGluValProLeuProGlyMetLysTrypVal	713
2593	AGGTGTAGT.....GATGGAACGAACCTGGA.....	2619
713	LaSPasnHisLysGlyValPheAsnValGluValAlaValAlaValSerSerI	730
2620	.GAGACSCCAAGTACCATAGTGGCCATTAACTCTTACTACGAAGAATGTC	2668
730	LeHisThrGlnAsp..ProTYrLeuAsp....LysPhePheAlaLeu	743
2669	TGTCAACAGCAAGATCTACTTTCTTTAGACATGCTCAAGTCTGTGGTTGGT	2718
744	...ValAsnAlaLeuAspGlnHisLysPheProValArgGlyLeuGlyAspMe	759
2719	TGTGTAGTACTAGCTGCTCAGACCAATCTGTGCTTTAGGCAAGCTGAAT	2768
759	TArgTLMetGluAsnAsnLeuGluLysGlnLysSerSerLysSerI	776
2769	TGGG.....AGGAATTTGTTAAATGT	2788
776	LaLeuAsnSerSerGlnLeuGluProValValArgPheLeuHisLeu...	791
2789	TAATGATCTCTGACAGCGTAGAAGCT...ACCAATTCCTGCACCTGGAT	2835
792LeuLeuAspLysLeu.....	796
2836	ATGTATCTAATGCTTTTAAAGAGCTCTCGAGAGAATACCCCTTCC	2885
797IleLeuLeuValIleArgProProValIleAlaGlyGlnI	810
2886	AATGGAAGATGTTCTTGAACCTTCTGAACCA.....CTATCCAAATG	2926
810	LeValAsnLeuGlyGlnAlaSerPheGluAlaMetAlaSerIleLeuAsn	826
2927	TGTGTTCTTTGTATGCTGCTGTGACCAAGATGTTGTAANAATATTTTAAAC	2976
827	Arg..LeuHis.....LysAsnLeu..GluGlyAsnHisAspGlnI	839
2977	CATGCTCTTCACTAGTATGATGTAACCAACCTAGCTCAACCAATATGAGCTGGA	3026
839	SGATyArgAsnSerLeuLeuAlaSerTYrIleHisTYrValPheArgLeuP	856
3027	GAACACAAAGGATGCTCAAGACAGATTTCTTACATAAT.....	3066
856	roAsnThrTYrProAsnSerSerSerProGlyProGlyLeuGlyGly	872
3067GGA	3069
873	SerValHisTYrAlaThrMetAlaArgSer.....AlaValArgPr	886
3070	GCAATTTTGGCATCTAGTCAAAAGAGAGAAATATATATTCTCTGTAAGAT	3119
886	oAlaSerLeuAsnLeuAsnArgSerValArgSerLeuSerAsnSerAsnPro.	902
3120	GGCCCTAGTAATATGTC.....CTTAAACATTGCTGAGGCTATGCTT	3163
903AspIleSerGlyThrProThrSerProasp	912
3164	ATTCAAAATGGGCCATCTTAATGTAAATGGGA...AAAGACTTTCCTGTGA	3210
913	AspGluValArgSer.....	918
3211	AATGAAGTATTTTACCAATTTCTTGTCGACATGATCAGCAAGTTGGAT	3260

5574 GAGAAATTCGCTTTCTTACACAAAGTTCCAGGATTTTTCACACAGCTGTCTTC 5622

1642 etIleuSerAspThrValLysMet..... 1650

5624 GACACTTCTCGCAAAAGACGACCGCATTCACAAACCCGCAAACTGGATTC 5673

1651 1656

5674 GAGTCAGAGCACTTTTTCGATGCTGTGATTAATAAATCAAAAGAAC 5723

1656 pProGluMetLeuIleAspLysMetLysArgLysGluLysGlyTyrGlnT 1673

5724 AATGCTTGCCTGTGTGCACTACATGACAGAGA.....CAAA 5788

1673 hrSerProGluArgLeuThr.....TyrLeuGlnAsn 1683

5759 AGAGACTTCTTCAGACAAATTTTAATGATGCTTTCGTGCTGGAT... 5805

1684 MetAlaGlyLysHisSerGluArgSerAsnHisAlaGluAlaGln.. 1699

5806TTAAATTCTACGAAGTGCACCAAGCT 5831

1700CysLeuValHisSerAlaIleValAlaGluTyrLeus 1713

5832 AGCTCACTCTTGCTGCTGCTGCTTACACTTACACTTACTC.....TATGCAG 5875

1713 erMetLeuGluAspArgLysTyrLeuPro..... 1722

5876 AAATCTATGCAGATTAAGAAAGTATGATGATCAACAGACAAAAGAGCTT 5925

1723ValGlyCysValThrPheGlnAsnIleSerSerAs 1734

5926 GCATTTGAGAAAGAGGAGGACGAGTACAACTATTCTAGCTGAGGAGAA 5975

1734 nValLeuGluGluSerAlaValSer..... 1742

5976 AAGTAAAGAGAAACGAGATTAAGTTACAGAGATCTCTTACAAATCT 6025

1743 ..AspAspValSerProAspGluGluGlyIleCysSerGly..... 1756

6026 ACGAAGTATAGGGGAGCCAGATAGTTGTATGCTGTGTGGAGGAG 6075

1757LysTyrPheThrGluSerGlyLeu 1765

6076 ATGTTCACAAACCATTCAGACTACGACACATATGAAACGAAACAGATGTG 6125

1765 IGIleu.....LeuGluGlnAla..AlaIleSer 1775

6126 GGGCAAAAGCCTAGTACATATGACTGCGAAACACAACTCCCTCATCA 6175

1775 heSerMetIleGlyMetTyrGluAlaValAsn..... 1785

6176 CACGCCAGGAGGAAATCTACAGCCCTTGCAAAATTTGGACTGCCAT 6225

1786GluValTyrLysValLeuIleProIleHisGluAlaAsnArgAs 1800

6226 ATTCCTTCGCTCAT.....TTAAAGAGTTGGATTGAAATAAAGA 6269

1800 p..AlaLysLysLeuSerThrIleHisGlyLysLeuGlnGluAlaPhe. 1815

6270 CTGGTCTCTCGAAGTACGAAGACTTCAT.....TACCACACACATGGA 6313

1816SerLysIleValHisGln 1821

6314 GGAATATGCAAGTGGGACCATTCGACTCCGTCAGCAAAAGAGTGAAGA 6363

1822 SerThrGlyTyrGluArgMetPheGlyThrTyr..... 1832

6364 ACCAGTTACCATGATCATGTATGCACATGCTCTACAAATCTCTAAGACAG 6413

1833PheArgValGlyPheTyrGlyThrLysPheGlyAspLeuAspGluG 1848

414 AGAATCTCTACATTTTATGAAGAAGCTCAAAATATGTCGCAAGTAAAGAG 6463


```

588 lnphemetYrGlyGluAspProSerAsnAlaMetProValIlePheGly 604
    |||
2342 AGCTATGT.....ACAGTTGCTTG..... 2361
605 LysSerSerCysserGluPheSerLysGluAlaYThrAlaValAlaY 621
    |||
2362 ...AGCACTGTCACAG..... 2376
621 rHisAsnArgSerProAsp.....PheHisGluGlu 632
    |||
2377 .....AAGAGTCCAAATAAGATTGCATCTGGCTTTTCTCCGATGTG 2419
632 lelyValLysLeuProAlaThrLeuThrAspHisHisLeuLeuPhe 648
    |||
2420 TAACTCAAAAGCTAATGATGACATTCGACATATTGTAAAGTTTGA 2469
649 ThrPheTyHisValSerCysGlnGlnLysGlnAsnThrProLeuGluThr 665
    |||
2470 TCCTTCATCAAAAAGCCATTGGACCGTGAGAGATAGATCAATCGAGA 2519
665 rProValGlyThrThrPrlPheProMetLeuGlnAsnGlyArgLeuLys 681
    |||
2520 TGATACT.....AATGAAATCTAATGC 2542
682 ....ThrGlyGlnPheCysLeuProValSerLeuGluLysProGln 696
2543 AGGTGAGAGATCATCCATCGATCTATTTAACGATTACCCGATAGT 2592
697 AlatySerValLeuSerProGluValProLeuProGlyMetLysTry 713
    |||
2593 AGTGTAT.....GATGCAACGACCTGGA..... 2619
713 LAspAsnHisLysGlyValPheAsnValGluValAlaValSerSer 730
    |||
2620 .GAGAGCCAAAGTACATAGCTGCCATTAACTTGTAGCTGAAGAATATC 2668
730 lEnHisThrGlnAsp...ProTyLeuAsp....LysPhePheAlaLeu 743
    |||
2669 TGTCAACCAAGATCTACTTTCTTAGACATGSCCAAGTTCTGTGTTG 2718
744 ...ValAsnAlaLeuAspGlnHisLeuPheProValArgIleGlyAsp 759
    |||
2719 TGTGTAACCTAGCTCAGACCAATACTGTCTTGTAGGCGACCTGAT 2768
759 lArgIleMetGluAsnAsnLeuGluAsnGlnLeuLysSerSerIleSer 776
    |||
2769 TCGG.....AGGAATTTGTAATGT 2788
776 lAlaLeuAsnSerSerGlnLeuGluProValValArgPheLeuHisLeu 791
    |||
2789 TAATGATTTCAAGCAGCTAGACST...ACCAATCCCTCCACCTGCAT 2835
792 .....LeuLeuAspLysLeu..... 796
2836 ATGTATCTAATGCTTTTAAAGAGCTCTCTGAGAGATACCCCTTGCC 2885
797 .....lIleLeuValIleArgProProValIleAlaGlyGln 810
2886 AATGAGAGATGTTCTTGAACCTTCGAACCA.....CTATCCAAATG 2926
810 lValAlaAsnLeuGlnAlaSerPheGluAlaMetAlaSerIleLeuAsn 826
    |||
2927 TGTGTTCTTTGATGCTGTCGACCAAGATGTTTGAATAATATTATAAC 2976
827 Arg...LeuHis.....LysAsnLeu...GluGlyAsnHisAspGlnHis 839
    |||
2977 CATGTCTTCATGATAGTGAATAAAGCTAGGTCAAAAGCAATATGAGACTCGA 3026
839 sGlyArgAsnSerLeuLeuAlaSerTyrlleHisTyrlValPheArgLeu 856
    |||
3027 GAACACAAAGGATGCTCAAGGACATTTCTTACAGTAAT..... 3066
856 roAsnThrTyProAsnSerSerProGlyProGlyGlyLeuGly 872
    |||
3067 .....CGA 3069
873 SerValHisTyAlaThrMetAlaArgSer.....AlaValArgPr 886
    |||
3070 GCATTTTGGCATCTTAACAAAGAGAGAAATATATATCTCTGTATAGAT 3119
886 oAlaSerLeuAsnLeuAsnArgSerArgSerLeuSerAsnSerPro 902
    |||
3120 GCCCTTACTAATTTGC.....CTTAAACTTTGCTTGAGCTGATCTT 3163
903 .....AspIleSerGlyThrProThrSerProAsp 912
    |||
3164 ATTCAAAATGGCCATTTCTTAATGTAATGGCA...AAGACTTTCCTGTA 3210
913 AspGluValArgSer.....l 918
    |||
3211 AATGAAGTATTTACAAATTTCTGCTGACAATCAATCACAAGATTGCAAT 3260
918 eIleGlySerLysGlyLeuAspArg..... 926
3261 GTTGGCTGCAGAGTCATATATAGATTGTTCCAGGACACAGAGGAGATT 3310
926 ..... 926
3311 CTTCAGGTTACTGAAGACCTTCTTGAACTTCAGCAACAGCTTTT 3360
927 SerAsnSerTyrlValAsnThrGlyProLysAlaAlaProTyrlGly 943
    |||
3361 GAAATGCATACTTG.....AAAGCTCAGGAAGAAATGAG 3395
943 rAsnProSerProSerAlaGluSerThrGlnAlaMetAspArgSerCys 960
    |||
3396 AGAATGTCCATAGTCGACGAAGCCCGAAGCTTGGATGAATTTATA 3445
960 snArgMetSerSerHisThrGluThrSerSerPheLeuGlnThrLeuThr 976
    |||
3446 ATAGA.....AAATCTGTTTACTGACGTTGATGA 3474
977 GlyArgLeuProHisLysLeuPheHisGluGlnLeuAlaLeuGlnTr 993
    |||
3475 GCTGTGTTTATCTCTGACCCCTATCTCGCAAAAACAGCTTTG...TT 3521
993 rValAlaCysSerGlySerValArgGluSerAlaLeuGlnGlnAlaTrp 1010
    |||
3522 TGCCCTGTGA...TCTGTGAAGAGAAATGATTAGAACTCAC... 3564
1010 hePhePheGluLeuMetValLysSerMetValHisHisLeuTyrlPheAsn 1026
    |||
3565 .....CTTGTGAAGAGGTTTACGAAAGTT... 3591
1027 AspLysLeuGlnAlaProArgLysSerArgPheProGluArgPheMetAs 1043
    |||
3592 .....TCTGAACCTTTGGATATAGAGCTTTAGA 3620
1043 rAspIleAlaAlaLeuValSerThrIleAlaSerAspIleValSerArg 1060
    |||
3621 AGACTT.....ATGGCATCTCATTTA... 3642
1060 heGlnLysAspThrGluMetValGluArgLeuAsn..... 1071
    |||
3643 .....GATTATCTGTGGTTTGAAGTGAATGCTTAATCTTCAAGTACTGAA 3684
1072 .....ThrSerLeuAlaPhePheLeuAsnAspLeuLeuSerValHe 1085
    |||
3685 TACAACCTTACTCTTTCTTCTTTATTTATTTAACTACACAAATATGTA 3734
1085 lAspArgGlyPheValPheSerLeuIleLysSerCysTyrlLys..... 1099
    |||
3735 GGAT.....TTCTATAGATCTTGTATTAAGGTTTGA 3766
1099 ..... 1099

```



```

3767 TTCACACTGCTGCTATTAGAGTCATTTTGATGAGGTGAAGTCATTGCT 3816
1099 ..... 1099
3817 AATCAGATTCAAGAGGACTGGAAAAAGCTCTCTAACAGACTGCTTCCAAA 3866
1099 ..... 1099
3867 GATTCGTGAATATATCTCTTATTTTGCCTATGAGGGTACACAGACACA 3916
1100 ..... GlnValSerSerLysLeuTyrSerLeuPro 1109
3917 GTGGAGTGCCACAGCAAGAGAGACTGCTACCAAGCTCTATGATGATGCTT 3966
1110 AsnProSerValLeuValSerLeuArgLeuAspPheLeuArgIleIecy 1126
3967 AAAAGTGAAGAACTATGTGGAAAAACAGATTGAT ..... 3999
1126 sSerHisGlnHisTyrValThrLeuAsnLeuPro ..... CysSerL 1140
4000 ..... CACTTATTCATTAGTAATTTACACAGATTTGTGTGGAGT 4039
1140 euleuThrProProAlaSerProSerProSerValSerSerAlaThrSer 1156
4040 TATTATATGACGTTACATGACCCAGCAAAATTTCTAGTCCAGTCAGAGCACT 4089
1157 GlnSerSerGlyPheSerThrAsnValGlnAspGlnLysIleAlaAsnMe 1173
4090 GACCTCTGTGACTTTTCAGGGGATTGGATCCTGCTCTAATCCACCACA 4139
1173 tPheGlnLeuSerValProPheArgGlnGlnHisTyrLeuAlaGlyLeuV 1190
4140 TTTTCCATCGCATGTGATTAAACAACTTTGCCATATATCAGCAATTTGTC 4189
1190 alLeuThrGlnLeuAlaValIleLeuAsp ..... ProAsp 1201
4190 ATAAACCAAGTTAAAGCATTTTAGAAATTTCTTCCAAAAGCCCTGAT 4239
1202 AlaGlnGlyLeuPheGlyLeuHisLysLysValIleAsnMetValHisAs 1218
4240 TCC ..... TATCAGAAATTT ..... 4254
1218 nLeuLeuSerSerHisAspSerAspProArgTyrSerAspProGlnIleL 1235
4235 ..... CTTCTTGCCATATGTGAGCAAGCAGCTGAACAAATATGTTATAGA 4303
1235 yAlaIArgValAlaMetLeuTyrLeuProLeuIleGlyIleIleMetGlu 1251
4304 ACCACAGAAATCTTAAATATATATACCTGTTGTTAGTTTATTACTGAAA 4353
1252 ThrValProGln ..... LeuTyrAsp.. 1258
4354 GATATTAAGAAAGTGGCTTAGAGAGACTTGGGCTTTGTTCTTCAGAGCT 4403
1259 ... PheThrGlnThrHisAsnGlnArgGlyArgProIleCysIleAlaT 1274
4404 TATTATATCTTTGATTCACTATATCAACCAAGGCCCTTGTATACATGG 4453
1274 hrAspAspTyrGlnSerGlnSer ..... GlySerMetIle 1285
4454 ATGTGTCATATAGTACGCTTCTCCCTTTGTTGTGACTTATTAGTCAGGTT 4503
1286 SerGlnThrValAlaMet ..... 1291
4504 TGCCAGACAGCCGCTGACTTACTGTAAAGATGCTGTAGAAAACATCTTCA 4553
1292 .AlaIleAlaGlyThrSerValProGlnLeuThrArgProGlySerPheL 1308
4554 TGTATATGTGTGTACATTATACCCCTTGTGTATGACGACGATGAGGTTTC 4603
1308 euleuThrSerThrSerGlyArgGlnHisThrThrPheSerAlaGlnSer 1324
4604 AGAACAGGTATGTGACTTGTGAAATACTTAGATGATTAACAGGAT 4653

```

```

1325 SerArgSerLeuLeuIleCysLeuLeuTrpValLeuLysAsnAlaAspG1 1341
4654 AATGAAAACCTCTATATCAGGATTAAGCTTTTAGATTCCTTTCTCCACCA 4703
1341 uThrValLeuGlnLysTrpPheThrAspLeuSerValLeuGlnLeuAsn. 1357
4704 TGTGTGT ..... TTTAAGGATTTGCCGTATTTACTACACAAAATAA 4741
1358 ..... ArgLeuLeuAspLeuLeu ..... Tyr 1364
4742 TCAATATACAGTACAGACCCCTTTTCACTCTTGAGAGAAATTAACATTTT 4791
1365 LeuCysValSerCysPheGlnuTyrLysGlyLysLysValPheGlnArgMe 1381
4792 CTCACACTAAGT ..... GTTATATATGACT 4817
1381 tAsnSerLeuThrPheLysLysSerLysAspMetArgAlaLysLeuGlu. 1397
4818 TCCATTGACAAAGACTTGAAAGACTAAAGATCTTCGAAGACACTGGAAC 4867
1398 ..... GlnAlaIleLeuGlySerIleGlyAlaArgGlnGlu ..... 1409
4868 TACATTAAGATACAGATGTGACATTATGAGACTTCTCAGGATAATTCGG 4917
1410 ..... MetValArgArgSerArgGlyGlnLeuGlnArgSerPr 1422
4918 CAAGATGGGATTTATGTGAAGACTAGTTGTCAATTTGTGTGACGATTACCA 4967
1422 oSer ..... Glys 1425
4968 GATGGCAATGAACACACACTGTGAAAAAGAAAGTTCTAGAGGCTGTGGAA 5017
1425 erAlaPheGlySerGlnGlnAsnLeuArgTrpArgLysAspMetThrHis 1441
5018 GCTGCTTGGGA ..... 5028
1442 TrpArgGlnAsnThrGlnLysLeuAspLysSerArgAlaGluIleGlnH 1458
5029 ..... GAAGTGGGTCCATATGATTTCTCTACCAATAGCTATACACA 5069
1458 sGlnAlaLeuIleAspGlyAsnLeuAlaThrGlnAlaAsnLeu ..... 1472
5070 TAGTAAA ..... GATGCATCTTATACCAAGGCCCTTAAGTTATTGAAG 5113
1473 ..... IleLeuAspThrLeuGlu ..... 1479
5114 ATTAAGACTTCAGTGCACCTTCATATGCTGACSTACSTGAATTAACACA 5163
1480 IleValAlaGlnThrValSerValThrGlnSerLysGlnSerIleLeuG1 1496
5164 CTGCTACAAAGATTTGTGCAAAGTTCGATCGACAGCTGTAACTGTTTGA 5213
1496 yGlyValLeuLysValLeu ..... LeuHisSerM 1506
5214 AAACATTTTGAAGCCAAAGACTGCACATAGTTTCTGAGCAATTTATAGA 5263
1506 etAlaCysAsnGlnSerAlaValTyrLeuGlnHisCysPheAlaThrGln 1522
5264 TGACAAACAGATCCCAATGCTGGCTATCTACAGCCT...TTTGAACAATCA 5310
1523 ArgAlaLeuValSerLysPheProGlnLeuLeuPheGlnGluGlnuThrG1 1539
5311 AGAAAAAAGTTTTTAGAGTACCAAGATTGACAAAGAAACCCCTTTTGA 5360
1539 uGlnCysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSer 1556
5361 AGGCTGTGATGATATA ..... 5376
1556 leglyThrIleArgSerHisProSerAlaSerLeuTyrLeuLeuMetArg 1572
5377 ..... AATCTGTGATCTCTTAAGT 5397

```

1573 GlnAsnPhgIuIleGly.....AsnAsp 1581
:::|||||
5398 GAAATTCATGACATTTGCATTAAGCACTGCTTGTCTTTTGGACAG 5447
1581 eAlaArgValIysMetGlnValPrometSerLeuSerLeuValGlyT 1598
:::|||||
5448 TGGAGGCAAAATGTGAATTCCTCAATTTAAAGCCAAATGTGTGAG 5497
1598 hrSerGlnAsnPhsAnGluIupheLeuArgSerLeuIleThrIle 1614
:::|||||
5498 TGAACCTGCATTT.....TGTCAAGCTGTA 5523
1615 LeuThrTyraIaGluIuAspLeuGluLeuArgGluThrThr 1628
||| |||
5524 CTTCCATCTTGTATCATGATATTTTACTCCAGATACAAATGAAATCG 5573
1629PheProAspGlnValGlnAspLeuValPheAsnLeuHisM 1642
:::|||||
5574 GAGAAATGCTTCTTACACATGTTCCAGGAAATTTTCCACAGCTGTCTC 5623
1642 eCileLeuSerAspThrValIysMet..... 1650
:::|||||
5624 GACACTTTCGCAACAGCCGATCCACACCCCTGCAAACTTGGATTC 5673
1651LysGluHisGlnGluAs 1656
5674 GAGTCAGAGCACTTTTCCGATGCTGTTGGATAAAAATCACAAGAC 5723
1656 pProGluMetLeuIleAspLeuMetTyraGlyIleAlaIysGlyGlnT 1673
:::|||||
5724 AATGCTTGTCTTGCATCATGAGAGA.....CAA 5758
1673 hrSerProGluArgLeuThr.....TripleGlnAsn 1683
||| |||
5758 AGAGACCTTCTTCAGAACAAATTTTAATGATGCTTCTGCTGGAT... 5805
1684 MetAlaGlyLysHisSerGluArgSerAsnHisAlaGluAlaGln.. 1699
||| |||
5806TTAAATTAATCTAGAAGTTGCCAAGT 5831
1700CysLeuValHisSerAlaAlaLeuValAlaGluTyLeuS 1713
||| |||
5832 ACCTGAGTCTTGTCTGCTCTCACTTACAGCTTACTC.....TATGAG 5875
1713 eMetLeuGluAspArgIysTyLeuPro..... 1722
:::|||||
5876 AAATCTATGCAATAGAAAAGTATGATGATCAAGAAAAAGATCTT 5925
1723ValGlyCysValThrPheGlnAsnIleSerSerAs 1734
5926 GCATTTGAAGAAGAACCCAGAGTCAACATTTCTACTGTGAGTGA 5975
1734 nValLeuGluIuSerAlaValSer..... 1742
||| |||
5976 AAGTAAAGAAAGAACTGCAATAGTTTACAGATCTTCTTAGAAATCT 6025
1743 ..AspAspValIaSerProAspGluGluGlyIleCysSerGly..... 1756
6026 ACAGAAGTATAGGAGGAGCCAGATAGTTGTATGCTGTGGAGGAG 6075
1757LysTyrPheThrGluSerGlyLeuVal 1765
6076 ATGTTCACAAACCATTACTAGACTAGCAACATATGAACAGCAATGTG 6125
1765 IolIyLeu.....LeuGluGlnAla..AlaIaSerP 1775
||| |||
6126 GGGCAAAAGCCCTAGTACATATGACCTGAAACAGCAATCCCTCATCA 6175
1775 heseMetAlaGlyMetTyrguAlaIaValAsn..... 1785
6176 CACGGCAGGAGGAATCATTCAGGCCCTTGCAAAATTTGGAGCTGCGCAT 6225
1786GluValTyIysValLeuIleProIleHisGluAlaAsnArgAs 1800

6226 ATTCCTTCCCTCAT.....TTAAAGATTTGATATGAAATATAAGA 6269
1800 P..AlaLysLysLeuSerThrIleHisGlyLysLeuGlnGluAlaPhe. 1815
6270 CTGGTGTCTGAAGTGAAGAACTTCAT.....TACCAAGCAGCATGGA 6313
1816SerLysIleValHisGln 1821
6314 GGAATATGCAAGTGGACCAATTCATTCCTCCGTCAGCAAGAAGTAGAAGA 6363
1822 SerThrGlyTrpGluArgMetPheGlyThrTy..... 1832
6364 ACCAGTTACCATGATCATCTGTACAAATGCTCTCAATCTCTTAGACAGAG 6413
1833PheArgValGlyPheTyrgIyThrLysPheGlyAspLeuAspGluG 1848
6414 AGAATTCCTACATTTTATGAAAGTCCAAATATGCCAGAGTAAAGAAAG 6463
1848 InGluPheValTyIysGlu.....Pro 1855
6464 TGGAAAGATGTGTAAAGCGACGCTTGAGTCTGTATTCGCTATGCC 6513
1856 AlaIleThrLysLeuAlaGluIleSerHisArgLeuGluGlyPheTyrgI 1872
6514 ACACTTAGCAGGTTTGCAGGCCATTTGA..GAGCTGGAAGAAGCATT...GG 6557
1872 yGluArgPheGlyGluAspValAl..... 1880
6558 GGAGCTTCTTCAGAAATGACGTCAACATGACAACTCTCTGAAGTATATA 6607
1881GluValIleLysAspSer..... 1886
6608 TTAAGTGGCAAAACACTCCACGCTTCTCAAGACAGTATTTTACTTTT 6657
1886 1886
6658 CAGGAGCCTATCATGGCTTACGCACAGCATTTTGGAGATCCTGATGGA 6707
1887AsnProValAspLysCysLysLeuAsp..ProAsnI 1898
6708 AAAGGAATGACAACTCACMAAGAGATGTATTAGGACATTTCTACCA 6757
1898 ySalArgIleGlnIleThrTyValGluProTyrrPheAspThrTyrgIu 1914
6758 AACACCTTAGAAGCTCTTACTGTGCGCAGAACTTCAAGAACACTGAG 6807
1915 MetLysAspArgIleThrTyrrPheAspLysAsnTyrrAsn 1927
6808 CTCCTGAAAGGGCAATATTTCAAAATTAACAGTACAAAT 6846
seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-08-952-127-2
seq_documentation_block:
; Sequence 2, Application US/08952127
; Patent No. 6211336
; GENERAL INFORMATION:
; APPLICANT: Shiloh, Yosef
; APPLICANT: Tagle, Danilo A.
; APPLICANT: Collins, Francis S.
; TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 6211336thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/952.127
: FILING DATE:
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: Kohn, Kenneth I.
: REGISTRATION NUMBER: 30,995
: REFERENCE/DOCKET NUMBER: 2290.00029
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 810-539-5050
: TELEFAX: 810-539-5055
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 9171 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: POSITION IN GENOME:
: CHROMOSOME/SEGMENT: 11q22-q23
: US-09-952-127-2

```

```

alignment_scores:
  Quality: 150.50      Length: 2063
  Ratio: 0.175         Gaps: 99
  Percent Similarity: 41.638   Percent Identity: 17.402

```

```

alignment_block:
US-09-737-246-2 x US-09-952-127-2 ..

```

```

Align seg 1/1 to: US-09-952-127-2 from: 1 to: 9171

```

```

267 ValLysCysLeuSerLeuLysPheGluIleGluIleGluProIlePheAl 283
    ||| |||||::||| ::::||| |||
1615 GTATGCTGTTGACTTTGGCACACGACCATGATATGTCAGGAACGCT 1664
283 aserLeuAlaLeuTyrAspValLysGluLysLysIleSerGluAsp 300
    :::::|||||::: |||:::|::: |||
1665 AAAAATGGGAAAT.....GAGCAAAATATGTGTGAATAAATA 1702
300 hetYrPheAsnLeuSerGluIleMetLysGlyLeu..... 313
    |||:::|||||:::|:::|:::|:::|:::|:::|:::|:::|
1703 GAAGCTTTCTTAAAGAAATCAATATGAAATGGCTTTATTCATCG 1752
314 .....ATGProHisValProProAlaAlaIleTh 323
1753 TTAGAGGTGACTTAGAAATAGCACAGAAAGTCCCTCA..... 1791
323 rThrIleAlaArgSerAlaIlePheSerIleThrTyrProSerGluAspV 340
    |||:::|:::|:::|:::|:::|:::|:::|:::|
1792 .....ATTCTTCACAGTAATTTCTCTCAT..... 1815
340 alPheLeuValIleLysLeuGluLysValLeuGlnGlnLysPheIleGly 356
    :::|:::|:::|:::|:::|:::|:::|:::|
1816 .....CTTGACTGGAGAAATCTTGAGTCTCACTAGTGA 1854
357 GluCysAlaGluProTyrMetIlePheLysGluAlaAspAlaThrLysAs 373
    :::|:::|:::|:::|:::|:::|:::|:::|
1855 AACTGTAAGCTGCATGAAATTTTCCAAAGGTCGCCGAATGTGAACA 1904
373 n.....LysGluLysLeuGluLysLeuSerGlnAlaAspGlnPheC 388
    :::|:::|:::|:::|:::|:::|:::|:::|
1905 CCACCAAAAGATTAAGAAAGAACTTCATTCACAGATGAAGAACTAT 1954
388 ysgLnArgLeuGlyLysTyrArgMetProPheAlaThrAlaIleHis 404
    :::|:::|:::|:::|:::|:::|:::|:::|
1955 TTCTTCAGCAACTTTTGACAAAGATGACTTT..... 1986
405 LeuMetAsnIleValSerSerAlaGlySerLeuGluArgAspSerThrGl 421

```

```

1987 ..TTAACCATTTGAGAGAAATGTGT...ATAGAAAGACACGATCCAG 2030
    :::::|||||::: :::::|:::|:::|:::|:::|:::|:::|
421 uValGluIleSerThrGlyLysGlySerTrpSerGluArgArg 438
    :::::|:::|:::|:::|:::|:::|:::|:::|
2031 TATTGGCTTCTGTGTCCACCAAGATCTC..... 2058
438 snSerSerIleValGlyArgArgSerLeuGluArgThrThrSerGlyAsp 454
    :::::|:::|:::|:::|:::|:::|:::|:::|
2059 .....AAGCAATCACTGATGCCGTCTCTTGCGA... 2088
455 AspAlaCysAsnLeuThrSerPheArgProAlaThrLeuThrValThrAs 471
2088 ..... 2088
471 nPhePheLysGlnGluGlyAspArgLeuSerAspGluAspLeuTyrLysP 488
    |||||:::|:::|:::|:::|:::|:::|:::|
2089 .....TTATCGAAACACCTTCTGAATATT 2113
488 heLeuAlaAspMetArgArgProSerSerValLeuArgArgLeuArgPro 504
    :::|:::|:::|:::|:::|:::|:::|:::|
2114 ACTCATCTGAGATTACAATTCAGAAACTGTGTCCGCTGTTCACGCTCT 2163
505 IleThrAlaGlnLeuLysIleAspIleSerProAlaProGluAsnProH 521
    :::::|:::|:::|:::|:::|:::|:::|:::|
2164 TTGGTGGGTGTCTTGCTGC..... 2184
521 sTyrCysLeuThrProGluLeuGlnValLysLeuTyrProAspSerA 538
    |||||:::|:::|:::|:::|:::|:::|:::|
2185 .TACGTTCACATGGGTGTAATAGCTGAAGAGAGACATATAAGCAAAAT 2233
538 rValArgProThrArgGluIleLeuGluPheProAlaArgAspValTyr 554
    :::|:::|:::|:::|:::|:::|:::|:::|
2234 TATTCCAAAAGCCACAGTCTTAATGCATGTGCAGAGAAAGTATC... 2280
555 ValProAsnThrThrTyrArgAsnLeuLeuTyrIleTyrProGlnSerI 571
    |||:::|:::|:::|:::|:::|:::|:::|
2281 .....ACTGTGTTAAAT.....AAGAC 2300
571 uAsnPheAlaAsnArgGlnGlySerAlaArgAsnIleThrValLysValG 588
    |||:::|:::|:::|:::|:::|:::|:::|
2301 AAATGAGAAATTCAGAAATGTGCTTGCAGAAATATG.....ATGC 2341
588 lnPheMetTyrGlyLysProSerAsnAlaMetProValIlePheGly 604
    ||:::|:::|:::|:::|:::|:::|:::|:::|
2342 AGCTATGT.....ACAGTTCCTG..... 2361
605 LysSerSerCysSerGluPheSerLysGluAlaTyrThrAlaValTyr 621
    ||:::|:::|:::|:::|:::|:::|:::|:::|
2362 ...AGCACTGTACCAAG..... 2376
621 rHisAsnArgSerProAsp.....PheHisGluGluI 632
    :::|:::|:::|:::|:::|:::|:::|:::|
2377 .....AAGAGTCCAAATAAGATTGCATCTGCGCTTTTCTCGGAATV 2419
632 leLysValLysLeuProAlaThrLeuThrAspHisHisIleLeuPhe 648
    |||||:::|:::|:::|:::|:::|:::|:::|
2420 TAACATCAAAAGCTAATGATGATGATGAGATTTGTAAGATTGCA 2469
649 ThrPheTyrHisValSerCysGlnGlnLysGlnAsnThrProLeuGluTh 665
    :::|:::|:::|:::|:::|:::|:::|:::|
2470 TCCTTCATCAAAAGCCATTTGACCGTGAGAGATAGAAATCAATGAAGA 2519
665 rProValGlyTyrThrTrpIleProMetLeuGlnAsnGlyArgLeuLys. 681
    :::|:::|:::|:::|:::|:::|:::|:::|
2520 TGATACT.....AATGGAATCTTAATGC 2542
682 .....ThrGlyLysPheCysLeuProValSerLeuGluLysProGln 696
    |||:::|:::|:::|:::|:::|:::|:::|
2543 AGGTGGAGATCAGTCATCCATGATCTATTTAAGCATTAACCTCGAATAGT 2592
697 AlaTyrSerValLeuSerProGluValProLeuProGluMetLysTrpVa 713
    :::|:::|:::|:::|:::|:::|:::|:::|

```



```

1173 tPheGluLeuSerValProPheArgGlnGlnHisTyrLeuAlaGlyLeuV 1190
    |||
    |||
    |||
1140 TTTTCATGCGCATGTGATTAAAGCAACATTTGCTATATCAACAATTGTC 4189
    |||
    |||
    |||
1190 aLeuThrGluLeuAlaValIleLeuAsp.....ProAsp 1201
    |||
    |||
    |||
4190 ATAAACCAAGTTTAAACGATTTAGAAATTCCTTCCAAAGCCCTGAT 4239
    |||
    |||
    |||
1202 AlaGlnGlyLeuPheGlyLeuHisTyrValIleAsnMetValHisAs 1218
    |||
    |||
    |||
4240 TCC.....TATCGAAATTT..... 4254
1218 nLeuLeuSerSerHisAspSerAspProArgTyrSerAspProGlnIleL 1235
    |||
    |||
    |||
4255 .CTTCTGGCATGTGTGACCAAGCAGCTGAAACAAATATGTTTATTAAG 4303
1235 yAlaArgValAlaMetLeuTyrLeuProLeuIleGlyIleIleMetGlu 1251
    |||
    |||
    |||
4304 AGCAGCAAAATCTTAAATATATACACCTGTGTATGTTTATTAAGAA 4353
1252 ThrValProGln.....LeuTyrAsp... 1258
    |||
    |||
    |||
4354 GATATATAAAGTGGCTTAGAGAGAGCTTGCCCTTGTCTTCGAGACGT 4403
1259 ....PheThrGluThrHisAsnGlnArgIleArgProIleCysIleAlaT 1274
    |||
    |||
    |||
4404 TATTATACCTTGTGATTCATATATCAACCAAGCCTCTGTATATCATG 4453
1274 hTrsAspArgTyrGluSerGluSer.....GlySerMetIle 1285
    |||
    |||
    |||
4454 ATGTGATTAATGATGAGTCTCCCTTGTGTGACTTATTAAGTCAGGTT 4503
1286 SerGlnThrValAlaMet..... 1291
    |||
    |||
    |||
4504 TGGCAGACAGCCCTGACTACTAGATGCTCTGAAACCATCTTCA 4553
1292 .AlaIleAlaGlyThrSerValProGlnLeuThrArgProGlySerPheL 1308
    |||
    |||
    |||
4554 TGTATATGTTGGTACACTATACCTGTGTATGACAGGTGAGTGCAGTC 4603
1308 eLeuLeuThrSerThrSerGlyArgGlnHisThrThrPheSerAlaGluSer 1324
    |||
    |||
    |||
4604 AGAAACAGTATTGACCTGTGAAATACCTAGTATAGATAACAAGGAT 4653
1325 SerArgSerLeuLeuIleCysLeuLeuTyrValLeuLysAsnAlaAspI 1341
    |||
    |||
    |||
4654 AATGAACACCTATATACAGATTAAGCTTTAGATCCCTTTCCTGACCA 4703
1341 uThrValLeuGlnLysTyrPheThrAspLeuSerValLeuGlnLeuAsn. 1357
    |||
    |||
    |||
4704 TGTGTGTT.....TTTAAGATTTGGCTATTACTCAGCAAAAAA 4741
1358 .....ArgLeuLeuAspLeuLeu.....Tyr 1364
    |||
    |||
    |||
4742 TCAAAATACAGTAGAGACCCCTTTTCACTCTGTGAGGAAATTAACATTT 4791
1365 LeuCysValSerCysPheGluTyrLysGlyLysValPheGluArg 1381
    |||
    |||
    |||
4792 CTCTCAGTAGT.....GTTATGATGCACCT 4817
1381 tAsnSerLeuThrPheLysLysSerLysAspMetArgAlaLysLeuGln. 1397
    |||
    |||
    |||
4818 TCCATTGACCAAGACTTGAAGAGCTTAAGGATCTTCAGACCAACTGAAAC 4867
1398 .....GluAlaIleLeuGlySerIleGlyValArgGlnGlu..... 1409
    |||
    |||
    |||
4868 TACATTAAGATCAAGATGGTGACATTAATGAGAGCTTCTCAGGATTAATCCG 4917
1410 .....MetValArgSerArgIleGlnLeuGlnLysArgSerP 1422
    |||
    |||
    |||
4918 CAAGATGGATTAATGTGTAAGAACTAATGTCATATTGTGACATTATCAAA 4967
1422 oSer.....Glys 1425

```

```

4968 GATGGCAATTAACCACTGGTGAAGAAAGATTCTAGAGCGTGTGGAA 5017
    |||
    |||
    |||
1425 eAlaPheGlySerGlnGluAsnLeuArgTyrArgLysAspMetThrHis 1441
    |||
    |||
    |||
5018 CTGCTGTGGG..... 5028
1442 TrpArgLysAsnThrGluLysLeuAspLysSerArgAlaGluIleGluH 1458
    |||
    |||
    |||
5029 .....GAGGTGGGTCTATAGATTTCTCCATCAGTATACAGCA 5069
1458 sGluAlaLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeu..... 1472
    |||
    |||
    |||
5070 TAGTAAA.....GATGCATCTTATACCAAGCCCTTAAGATTATTTGAG 5113
1473 .....IleIleLeuAspThrLeuGlu..... 1479
    |||
    |||
    |||
5114 ATAAAGAACTTCAGTGCACCTTCATTAAGTGTGACCTTACGATTAACACA 5163
1480 lLeuValAlaGlnThrValSerValThrGluSerLysGluSerIleLeuG 1496
    |||
    |||
    |||
5164 CTGTGAGAGATTTGTCTCAAGTTGCATACAGCAGCTGTACCTGTTGAA 5213
1496 yGlyValLeuLysValLeu.....LeuHisSerM 1506
    |||
    |||
    |||
5214 AAACATTTTACCCACCAAGACTGACATAGTTCTGGAGAGATTATTAAGA 5263
1506 eAlaCysAsnGlnSerAlaValTyrLeuGlnHisCysPheAlaThrGln 1522
    |||
    |||
    |||
5264 TGACACAGATCCAAATGCTGGCCTTATCTACAGCCT...TTTGAACAATCA 5310
1523 ArgAlaLeuValSerLysPheProGluLeuLeuPheGlnGluThrGln 1539
    |||
    |||
    |||
5311 AGAAAAAGTTTATAGAGTACCCAGATTTGACAAGAAACCCCTTTTGA 5360
1539 uGlnCysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSer 1556
    |||
    |||
    |||
5361 AGGCCGTGATGATTA..... 5376
1556 lGlyThrIleArgSerHisProSerAlaSerLeuTyrLeuLeuMetArg 1572
    |||
    |||
    |||
5377 .....AATCTGTGATTCCTTAAGT 5397
1573 GlnAsnPheGluIleGly.....AsnAsnPh 1581
    |||
    |||
    |||
5398 GAAATTCATGACATTTGGATTAAGCACTGACTGTGCTTTTGGACAG 5447
1581 eAlaArgValLysMetGlnValPrometSerLeuSerSerLeuValGlyT 1598
    |||
    |||
    |||
5448 TGGAGGCACAAATGTGAATTCCTCAATTTTAAAGCCAATGTGTGAG 5497
1598 hTrsSerGlnAsnPheAsnGlnGluPheLeuArgArgSerLeuLysThrIle 1614
    |||
    |||
    |||
5498 TGAATACGTGACTT.....TGTCAAGACTGTA 5523
1615 LeuThrTyrAlaGluGluAspLeuGluLeuArgGluThrThr..... 1628
    |||
    |||
    |||
5524 CTTCATATCTGATTCATGATATTTACTCCAAGATTAACAAATGAATCATG 5573
1629 .....PheProAspGlnValGlnAspLeuValPheAsnLeuHisM 1642
    |||
    |||
    |||
1642 eTlleLeuSerAspThrValLysMet..... 1650
    |||
    |||
    |||
5624 GACACTTCTCGCAACGAGCCGATCCACACCCCTGCAAACTTGATTTCA 5673
1651 .....LysGlnHisGlnGluAs 1656
    |||
    |||
    |||
5674 GAGTCAGAGCACTTTTCCGATGCTGTTTGATTAATAAATACAAAGAAC 5723
1656 pProGluMetLeuIleAspLeuMetTyrArgIleAlaLysGlyTyrGlnT 1673
    |||
    |||
    |||

```


alignment block:
US-09-737-246-2 x US-08-952-014C-2 ..

Align seg 1/1 to: US-08-952-014C-2 from: 1 to: 9171

```

267 ValIysCysLeuSerLeuLysPheGluIleGluIleGluProIlePheal 283
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1615 GTATGCTGTTGACTTTGGCACTGACACCAAGATATGTTCCAGGAAGCT 1664
283 aserLeuAlaLeuTYrAspValIysGluLysLysIleSerGluAsn 300
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1665 AAAAATGGGAATA.....GAGCAAAATATGTGTGAATGAATA 1702
300 hetYrPheAspLeuAsnSerGluIleMetIysGluLeu..... 313
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1703 GAAGCTTTCTTAAGCATCATATGAATGCGCTCTTATTCATCAG 1752
314 .....ArgProHisValProProAlaAlaIleTh 323
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1753 TTAGAGGGTGAAGTAAATAGCACAGAAAGTCCCA..... 1791
323 rthIleuAlaArgSerAlaIlePheSerIleThTYrProSerGluAspV 340
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1792 .....ATTCTTCACAGTAAATTTCTCTCAT..... 1815
340 alPheLeuValIleLysLeuGluLysValIleGluIleGluLysPileGly 356
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1816 .....CTGTACTGAGAAATATTTCTGTGAGTCTCAGTGAATA 1854
357 GluCysAlaGluProTYrMetIlePheLysGluAlaAspAlaThrLysAs 373
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1855 AACTGTAAGCTGCAATGAAATTTTCCAAAGCGTCCAGAAATGGAACA 1904
373 n.....LysGluLysLeuGluLysLeuLysSerGlnAlaAspGlnPheC 388
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1905 CCACCAAAAAGATAAAGAAAGAACTTCATTCAGAAAGTGAAGAACTAT 1954
388 ysgIleArgLeuGlyLysTYrArgMetProPheAlaThrPheAlaIleHis 404
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1955 TTCTTGACACACTTTTGACAAAGATGACTT..... 1986
405 LeuMetAsnIleValSerSerAlaGlySerLeuGluArgAspSerThG1 421
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1987 ...TTAACCATTTGTGAGAAATGTGT...ATAGAAAACACCAAGTCCAG 2030
421 uValGluIleSerThrGlyLysArgLysLysSerThrSerGluArgArgA 438
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2031 TATTGGCTTCTGTCCACCAAGATCTC..... 2058
438 snSerSerIleValIysArgArgSerLeuGluArgThrThrSerGlyAsp 454
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2059 .....AAGGATCAGCTGATCGCTGTCTGTGGA... 2088
455 AspAlaCysAsnLeuThrSerPheArgProAlaThrLeuThValThras 471
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2088 ..... 2088
471 nPhePheLysGluGluLysArgArgSerAspGluAspLeuTYrLysP 488
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2089 .....TTATCAGAAACAGTTCTGAATTAAT 2113
488 heLeuAlaAspMetArgArgProSerSerValLeuArgArgLeuArgPro 504
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2114 ACATCATCTGATTAACAATTCAGAAACCTTGTCCGGTTCACAGCTTT 2163
505 IleThraIleGluLysIleAspIleSerProAlaProGluAsnProH1 521
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2164 TTGGTGGTGTCTGTGCTGC..... 2184
521 sTYrCysLeuThrProGluLeuGluValLysLeuTYrProAspSerA 538
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2185 .TACTGTACATGGGTGTAAATAGCTGAAAGAGGAAATATAGTCAAGAT 2233

```

```

538 rGValArgProThrArgGluIleLeuGluPheProAlaArgAspValTYr 554
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2234 TATTCCAGAAAGCAAGTCTCTATGCAATGTGCAGAGAAATATC... 2280
555 ValProAsnThrThTYrArgAsnLeuLeuTYrIleTYrProGlnSerIle 571
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2281 .....ACTCGTTTAAAAAT.....AAGAC 2300
571 uAsnPheAlaAsnArgGlnGlySerAlaArgAsnIleThValLysValG 588
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2301 AAATGAGCAATTCAGAAATGTGCTTGTGAGAAATATG.....ATGC 2341
588 LnPheMetTYrGlyLysProSerAsnAlaMetProValIlePheGly 604
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2342 AGCTATGT.....ACAGTGTCTG..... 2361
605 LysSerSerCysSerGluPheSerLysGluAlaTYrThraValValTY 621
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2362 ...AGCAACTGTACCAAG..... 2376
621 rHisAsnArgSerProAsp.....PheHisGluGluI 632
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2377 .....AAGAGTCCAAATTAAGATTGCATCGCTTTTCTCGGATGT 2419
632 IeLysValLysLeuProAlaThrLeuThraPheHisHisLysLeuPhe 648
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2420 TAACATCAAAAGCTAATGATGACATTCAGATATTGTAAGATTGCA 2469
649 ThrPheTYrHisValSerCysGlnGlnLysGlnAsnThrProLeuGluTh 665
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2470 TCCTTCATCAAAAAGCAATTGACCGTGAGAGTACATCATCAATGAGA 2519
665 rProValGlyTYrThrTrpIleProMetLeuGlnAsnGlyArgLeuLys 681
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2520 TGATACT.....AATGCAATGTAATG 2542
682 .....ThrGlyLysPheCysLeuProValSerLeuGluLysProProGln 696
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2543 AGGTGAGAGATCAGTCCATGATCATGATTAATGACGTACCGTGAATG 2592
697 AlaTYrSerValLeuSerProGluValProLeuProGlyMetLysTrpVa 713
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2593 AGGTGATGT.....GATGCAACGACACCTGCA..... 2619
713 LAspAsnHisLysGlyValPheAsnValGluValAlaValAspSerI 730
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2620 .GAGAGCCAAAGTACACATAGTGGCATTAATCTTTAGCTGAGAAATATC 2668
730 IeHisThrGlnAsp...ProTYrLeuAsp.....LysPhePheAlaLeu 743
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2669 TGTCAAAAGCAAGATCTACTTTTCTTAGACATGTCAGATCTGTGTGTTG 2718
744 ...ValAsnAlaLeuAspGlnHisLeuPheProValArgIleGlyAspMe 759
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2719 TGTGTAACTACTGCTCAGACCAATATCTGTGCTTGTAGGCGCATGAT 2768
759 LArgIleMetGluAsnAsnLeuGluAsnGluLeuLysSerSerIleSera 776
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2769 TCGG.....AGCAATGTTAATGT 2788
776 lAlLeuAsnSerSerGlnLeuGluProValValArgPheLeuHisLeu... 791
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2789 TAATGATTTAGACAGCTAGAACST...ACCAATCCSTCCACSTGCAT 2835
792 .....LeuLeuAspLysLeu..... 796
2836 ATGTATCTAATGCTTTTAAAGAGCTTCTGTGAGAGAGTACCCCTTGCC 2885
797 .....lIleLeuLeuValIleArgProProValIleAlaGlyGlnI 810
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2886 AATGAGAGATGTCTTGAACCTTGAACA...STATCCAAATG 2926
810 lValAlaLeuGlyGlnAlaSerPheGluAlaMetAlaSerIleIleAsn 826

```

3643GATTATCTCGGTTTGGATAGGCTAAATCTTCAGATACGAA	3664
1072ThSerLeuAlaPhePheLeuAsnAspLeuSerValMe	1085
3685	TACAACCTAACCTCTTCTTCCCTTTATTTATTTAACTACACAAATATGGA	3734
1085	TAspATrgGLyPheValPheSerLeuIleYsSerCysTrpLys.....	1099
3735	GGAT.....TTCTATAGATCTTGTTATAGCGTTTGA	3766
1099	1099
3767	TTCCACATCGTGATGATTAAAGTCAATTTGATGAGGTAGGCCATTGCT	3816
1099	1099
3817	AATCAGATTCAGAGAGACTGGAAAAAGTCTCTAACAGACTGCTTCCAA	3866
1099	1099
3867	GATTCTTGTAANAATTCCTTATTTTGGCATAGAGGTACAGAGACA	3916
1100GlnValSerSerLysLeuTrpSerLeuPro	1109
3917	GTGGAGTGCACACAGCAAAAGAGACTGCTACCAAGCTATGATATGCTT	3966
1110	AsnProSerValLeuValSerLeuTrpGluSerPheLeuArgIleIleCys	1126
3967	AAAAGGMAAACTTATTTGGAAAAACAGATTGAT.....	3999
1126	sSerHisGLuHisTrpValThrLeuAsnLeuPro.....CysSerL	1140
4000CACTATTCTATTAGTAATTTTACAGAGACTGTGGTGAGGT	4039
1140	eueuthrProProAlaSerProSerProSerValSerSerAlaThrSer	1156
4040	TATTGTAGACGTTACATGAGCCAGCAAAATTCAGTGCACGACAGACT	4089
1157	GlnSerSerGlyPheSerThrAsnValGlnAspGlnLysIleAlaAsnMe	1173
4090	GACCTCTGACATCTTTACAGGAGATTGGATGCTGCTCAATTCACCTCA	4139
1173	thPheGluLeuSerValProPheArgGlnGlnHisTrpLeuAlaGlyLeu	1190
4140	TTTTTCATGCGCATGTGATTAAGAACAACATTTGCCATATACGCAATTGTC	4189
1190	alLeuThrGluLeuAlaValIleLeuAsp.....ProAsp	1201
4190	ATPAAACCAAGTTAAAGCAATTTAGAAATCTTCCAAAGGCCGAT	4239
1202	AlaGluGlyLeuPheGlyLeuHisLysLysValIleAsnMetValHisAs	1218
4240	TCC.....TATCAGAAATTT.....	4254
1218	nleuLeuSerSerHisAspSerAspProArgTrpSerProGlnIleL	1235
4255	CTCTGTGCGCATGTGAGCAACAGCGTGAACAATAATGTTTATAGCA	4303
1235	ysAlaArgValAlaMetLeuTrpLeuProLeuIleGlyIleIleMetGlu	1251
4304	AGCAGCAAGATTCCTTAAATATATATACACCGTGTGTAGTTATATACGAA	4353
1252	ThrValProGln.....LeuTrpAsp.....	1258
4354	GATATPAAAAAGTGGCTTAGGAGAGACTTTGGCCCTTTGCTTCGAGACGT	4403
1259PheThrGluThrHisAsnGlnArgGlyArgProIleCysIleAlaT	1274
4404	TATTTATACTTGTGATTCACATATACACCAAAAGCCCTTGTGATCATGG	4453
1274	hrAspAspTrpGlyGluSerGluSer.....GlySerMetIle	1285
4454	ATGTGTCACTTACAGTACCTTCCCTTTGTGTGTGACTATTATTAAGCAAGTT	4503


```

1286 SerGlnThrValAlaMet..... 1291
1287 |||||.....
4504 TCCGACAGACGCGGTACTTACTGAAGATGCTAGAAAAACATCTTCA 4553
1292 .AlaIleAlaGlyThrSerValProGlnLeuThrArgProGlySerPheL 1308
1293 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4554 TGTATTTGTTGGTACATTATACCCCTTGATGACGAGGTTC 4603
1308 euleThrSerThrSerGlyArgGlnHisThrThrPheSerAlaGlySer 1324
1309 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4604 AGAAACAGATTGGACTTGTGAATCTAGTATGATAGATACAGAGAT 4653
1325 SerArgSerLeuLeuIleCysLeuLeuTrpValLeuLysAsnAlaAspG1 1341
1326 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4654 AATGAAACCTCTATATACAGATTAGCTTTAGATCCTTTCTCGACCA 4703
1341 uThrValLeuGlnLysTrpPheThrAspLeuSerValLeuGlnLeuAsn. 1357
1342 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4704 TGTGTGT.....TTAAGATTTCGATTTACTCAGCAAAAAA 4741
1358 .....ArgLeuLeuAspLeuLeu.....Tyr 1364
1359 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4742 TCAAAATACAGTAGAGACCTTTTTCACCTCTGGAGAAATTACCATTTT 4791
1365 LeucysValSerCysPheGluLysGlyLysValPheGluArgme 1381
1366 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4792 CTCACGATAGT.....GTTTATGATGCACT 4817
1381 tAsnSerLeuThrPheLysLysSerLysAspMetArgAlaLysLeuGln. 1397
1382 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4818 TCCATTGACAGACTTGAAGACTTGAAGATCTTGAAGACACACTGGAAC 4867
1398 .....GluAlaIleLeuGlySerIleGlyAlaArgGlnLys..... 1409
1399 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4868 TACATTAAGATCAGATGGTGGACATTATGAGAGCTTCAGAGATTAATCG 4917
1410 .....MetValArgArgSerArgGlyGlnLeuGluArgSerPr 1422
1411 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4918 CAAGATGGGATTATGTGTAACTACTTGTCAATTGTTGCAAGTATCCAA 4967
1422 oSer.....Glys 1425
1423 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4968 GATGGCAATTAACACACTGTGTAAGAAAGATTCTAGAGCTGTGGAA 5017
1425 eAlaIlePheGlySerGlnLysLeuArgTyrPheGlyAspMetThrHis 1441
1426 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5018 GCTGCTGGGA..... 5028
1442 TrpArgGlnAsnThrGluLysLeuAspLysSerArgAlaGluIleGluH1 1458
1443 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5029 .....GAAGTGGTCCATATAGATTCTCTACCATAGCTATACACA 5069
1458 sGluAlaLeuIleAspGlyAsnLeuAlaThrGluAlaAsnLeu..... 1472
1459 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5070 TAGTAAA.....GATGCATCTTATACCAAGCCCTTAAGTTAATTGAAG 5113
1473 .....IleIleLeuAspThrLeuGln..... 1479
1474 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5114 ATTAAGAACTTCAGTGAACCTTCATATAGCTGACCTACCTGAAATTAACA 5163
1480 lIleValIleGlnThrValSerValThrGluSerLysGluSerIleLeuG1 1496
1481 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5164 CTGGTAGAAGATTGTGTCAAGTTGATGACAGACGCTTACCTGTTTGA 5213
1496 yGlyValLeuLysValLeu.....LeuHisSerW 1506
1497 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5214 AAACATTTTATGCCCAAGACTGACATAGTTTCTGGAGCATTTATAGA 5263
1506 eAlaIleCysAsnGlnSerAlaValTyrLeuGlnHisCysPheAlaThrGln 1522
1507 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5264 TGACAAACGATCCATATGCTGGCTATCTACAGCT...TTTAAACAATCA 5310

```

```

1523 ArgAlaLeuValSerLysPheProGluLeuLeuPheGluGluThrG1 1539
1524 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5311 AGAAAAAGTTTTTAGAAGTACCACAGATTTGACAAAGAAACCCCTTTGA 5360
1539 uGlnCysAlaAspLeuCysLeuArgLeuLeuArgHisCysSerSerSeri 1556
1540 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5361 AGGCTGAGATGATATA..... 5376
1556 IeGlyThrIleArgSerHisProSerAlaSerLeuTyrLeuLeuMetArg 1572
1557 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5377 .....AATCTGTGATTCCTTAAGT 5397
1573 GlnAsnPheGluIleGly.....AsnAsnPh 1581
1574 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5398 GAAATATATACATTTTGGATTAAGACACAGCTGCTGCTTTTGGACAG 5447
1581 eAlaArgValLysMetGlnValProMetSerLeuSerSerLeuValGlyT 1598
1582 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5448 TCGAGGCACAAAATGTAATCTTCAATATTAAGCCAAATGTGGAAG 5497
1598 hrSerGlnAsnPheAsnGluLysPheLeuArgArgSerLeuLysThrIle 1614
1599 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5498 TGAACACTGACTTT.....TCTCAGACTGTA 5523
1615 LeuThrTyrAlaGluGluAspLeuGluLeuArgLysThr..... 1628
1616 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5524 CTTCATACACTTGATTCATGATATTTTACTCCAGATACAAATTAATCAG 5573
1629 .....PheProAspGlnValGlnAspLeuValPheAsnLeuHisM 1642
1630 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5574 GAGAAATCTGCTTTCTACACATGTTCAAGGATTTTTCACCAGCTGTTC 5623
1642 eTlleLeuSerAspThrValLysMet..... 1650
1643 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5624 GACACTTCTGCAACAGACCGATCCACAAACCCCTGCAACTTGATTA 5673
1651 .....LysGluHisGlnGlnLys 1656
1652 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5674 GAGTCNAGACACTTTTCCGATGCTGTTGGATTAATAAATCACAAGAAC 5723
1656 ProGluMetLeuIleAspLeuMetTyrArgIleAlaLysGlyTyrGlnT 1673
1657 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5724 AATGCTTGTGTTGTGACATACATGAGAGAA.....CAAA 5758
1673 hrSerProGluArgLeuThr.....TrpLeuGlnAsn 1683
1674 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5759 AGACACCTTCTTCAAGAACATTTTAATGATGCTTTCTGGCTGGAT... 5805
1684 MetAlaGlyLysHisSerGluArgSerAsnHisAlaGluAlaIleGln.. 1699
1685 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5806 .....TAAATTTATCTGAAGTCCCAAGT 5831
1700 .....CysLeuValHisSerAlaIleValAlaGluTyrLeus 1713
1701 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5832 AGCTCAGTCTGTGCTGCTCCTTACAGCTTATAC.....TATGACAG 5875
1713 eMetLeuGluAspArgLysTyrLeuPro..... 1722
1714 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5876 AAATCTATGACATTAAGAAAGTATGATGATGATCAAGAAAAAGACTCTT 5925
1723 .....ValGlyCysValThrPheGlnAsnIleSerSerAs 1734
1724 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5926 GCATTTAAAGAAAGAACCCAGATACACATATTTCTTACCTTGAGTAAAA 5975
1734 nValLeuGluLysSerAlaValSer..... 1742
1735 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5976 AAGTAAGAAGAAACTGGAATAGTTTACAGATCTTCTTGAATCTA 6025
1743 .AspAspValIleSerProAspGluGluGlyIleLysSerGly..... 1756
1744 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6026 ACAGAAATATAGGGAGCCAGATAGTTTGTATGCGTGTGGAGGGAAG 6075
1757 .....LysTyrPheThrGluSerGlyLeuVal 1765

```



```

150 .....GlnLysGlyLeuProLysGlnVal.....PheGluSe 160
    |||::||| |||::||| |||::|||
4114 CTGCATGCAACAGGAACTGTGAAAAATCTTTAGAGAGGTTCTTGAG 4163
160 RASpGlnAlaProAsnSpGlnAsnSerTyrGlnAspAspGlnAspLeuL 177
    |||::||| |||::||| |||::|||
4164 TATTTTAACACAGAGATGATTCGATTTTCA..... 4194
177 YsArgSerMetSerLeIleAspSphrProArgGlySerTrpAlaCys 193
    |||::||| |||::||| |||::|||
4195 ..... 4200
194 SerIlePheAspLeuLysAsnSerLeuProAspAlaLeuLeuProAsnLe 210
    |||::||| |||::||| |||::|||
4201 .....CAGAGAGTATTTGGTGATCTTTGGTATCTTGAT 4235
210 uLeuAspArgThrProAsnGlnGluIleLeuAspArgGlnAsnAspSpGlnA 227
    |||::||| |||::||| |||::|||
4236 GAGTTCAAGAACATGTTCAAGAGCT.....AACCTTCTTTTG 4275
227 rGlySerSerAsnArgHisLysGlnLeuPheAlaLeuHisProSerProAsp 243
    |||::||| |||::||| |||::|||
4276 AGAATTTTCTGGAGAAATCTCTTGACAAAATTTCTTCTTGCGGTAT 4325
244 GluGlu.....GluProIleGluArgLeuSe 252
    |||::||| |||::||| |||::|||
4326 TCTGAAATATATTGAAGTGAATGACTACTAGACCTTCACAGATCTAA 4375
252 rValProAspIleProLysGlnHisPheGlyGlnArgLeuLeuValLysC 269
    |||::||| |||::||| |||::|||
4376 CTTTCCTTTACTGCACGCTCCAAATTTAAAGCAAGGTTTTCATCACA 4425
269 yLeuSerLeuLysPheGlu.....IleGluIleGlu 279
    |||::||| |||::||| |||::|||
4426 AAGTATCTCGGATTTTAAACAGTAGAGCCATGGTTATTGACAAAGAGC 4475
280 ProIlePheAlaSerLeuAlaLeuTyrAspValLysGluLysLysI 296
    |||::||| |||::||| |||::|||
4476 ACGAGTTTC..... 4484
296 eSerGluAsnPheTyrPheAspLeuAsnSerGlnIleMetLysGlyLeuL 313
    |||::||| |||::||| |||::|||
4485 .....ACGAGCAAGAAAGAGCGTCAT 4506
313 euArgProHisValProProAlaAlaIleThrThrLeuAlaArgSerAla 329
    |||::||| |||::||| |||::|||
4507 AGAGAGATTTTCCCATCGCGCTGCTTC...ATCTTGCCACATAGC... 4550
330 IlePheSerIleThrTyrProSerGlnAspValPheLeuValIleLysLe 346
    |||::||| |||::||| |||::|||
4551 .....CCCATGTCACCTGCCGTTGCTGGGCAAAACT 4582
346 uGluLysVal..LeuGlnGlnGly.....AspIle 355
    |||::||| |||::||| |||::|||
4583 GCTGGCCACACTATCAAGAGGTTTCACTGTTCCCTGGTTAAATGTC 4632
356 GlyLysCysAlaGluProTyrMetIlePheLysGlnAlaAspAlaThrLy 372
    |||::||| |||::||| |||::|||
4633 ...GAGTGT.....ATCCATGAGCTGAGAGACTAC 4661
372 sAsnLysGluLysLeuGluLysLeuLysSerGlnAlaAspGlnPheCysG 389
    |||::||| |||::||| |||::|||
4662 AGAAAAAGAAAGAGATTAAGAAAGAAACAATCA..... 4698
389 InArgLeuGluLysTyrArgMetProPheAlaIleThrAlaIleHisLeu 405
    |||::||| |||::||| |||::|||
4699 .....TTA 4701
406 MetAsnIleValSerSerAlaGlySerLeuGluArgAspSerThrGlu.. 421
    |||::||| |||::||| |||::|||
4702 ATTTTACCGATAGCAGTTTGATGTACAGAGACGACAGACCAAGAG 4751
422 ...ValGluIleSerThrGlyGluArg..... 429

```

```

4752 TGCAGAGTACATAAATCTGTGTGAAGACTATAGAGAAGAGTATTC 4801
    |||::||| |||::||| |||::|||
430 .....LysGlySer.....TrpSerGlu 435
    |||::||| |||::||| |||::|||
4802 ATATTAATTTCACTGGATCCAAAGCCTTGATGATCCAACTGGGCTGAT 4851
436 ArgArgAsnSerSerIleValGlyArgArgSerLeuGluArgThrThSe 452
    |||::||| |||::||| |||::|||
4852 CCCACAAATGCCACTCTTATCTTGCTGTGTGCATGAT.....TC 4892
452 rGlyAspAsp.....AlaCysAsnLeuThrSer..... 461
    |||::||| |||::||| |||::|||
4893 AAATGATGACATGAAGACTGTTTACTACACAGCTGGAATCAGAGAGA 4942
462 ....PheArgProAlaThr.....LeuThrValThrAsnPhePhe 473
    |||::||| |||::||| |||::|||
4943 ATATTTTCTCCAGACAAATGCGAAACATTTAGTACTCACTACTTACAG 4992
474 LysGlnGluGlyAspArgLeuSerAspGluAspLeuTyrLysPheLeuAl 490
    |||::||| |||::||| |||::|||
4993 CACCCCAAGGAAAGAGAGATTCATGGGAAATCTCCATATGGGTCTC 5042
490 aAspMetArgArgProSerSerValLeuArgArgLeuArgProIleThrA 507
    |||::||| |||::||| |||::|||
5043 TGGACAGAGGAGACCTGATGTTACTTTGATTTTATCTTCCAGAAAAA 5092
507 IagIleLysLysIleAspIleSerProAlaProGluAsnProHisTyrCys 523
    |||::||| |||::||| |||::|||
5093 CAAGTTGTATCTGAT.....AGCATATAAACATTTTGC 5127
524 LeuThrProGluLeuLeuGlnValLysLeuTyrProAspSerArgValAr 540
    |||::||| |||::||| |||::|||
5128 ATGATGGCCATGTTTATCATCCCA..... 5154
540 gProThrArgGluIleLeuGlnPheProAlaArgAspValTyrValProA 557
    |||::||| |||::||| |||::|||
5155 .....GAAGAGTTTTCGAGTTGGCTGGGAAAA..... 5181
557 snThrThrTyrArgAsnLeuLeuTyrIleTyrProGlnSerLeuAsnPhe 573
    |||::||| |||::||| |||::|||
5182 ..TGGACCTGGGAATTTGCTCTCTTC..... 5208
574 AlaAsnArgGlnLysSerAlaArgAsnIleThrValLysValGlnPheMe 590
    |||::||| |||::||| |||::|||
5209 .....AAGGAGCTAAGTTGGTTCACAGAGGCCCTTTATCT 5246
590 tTyr.....GlyGluAspProSerAsnAlaMetProValIlePheGlyL 605
    |||::||| |||::||| |||::|||
5247 GTATGCTTGTGGACCCACCATCATCTGTAATGTCATGATATGAGCA 5296
605 ySerSerCysSerGluPheSerLys..... 613
    |||::||| |||::||| |||::|||
5297 AG...CCAGTCAATGACTACTCCAAATATTAATTAAGAAATTTTGGCA 5343
614 .....GluAlaTyr..... 616
    |||::||| |||::||| |||::|||
5344 TGTGAACAATTCAGAGACTTTTATGACCACAGAAAGATGATTTGG 5393
617 .....ThrAlaValValTyrHisAsnArgSerProAsp 627
    |||::||| |||::||| |||::|||
5394 TCTCTTAATTAAGAAAGCTTTCAGTTGTTTATACAACTTACTGCTCTC 5443
628 .....PheHisGluGluIleLysValLys..... 635
    |||::||| |||::||| |||::|||
5444 AGTATACCATCTATGAACAGATTAAGACTTAAGATCAAAATGAAGAAC 5493
636 .....LeuProAlaThrLeuTh 641
    |||::||| |||::||| |||::|||
5494 CAACCTCTCAAGACCTTCACCTCAAAAGAGTTTCAGAGCATCTTATTT 5543
641 RASpHisHisLeu..... 646
    |||::||| |||::||| |||::|||

```

5544 AGAAGCTCATCATATAAAAAATCTCAACCTACTGATATATAAACTATTC 5593
 647 Leuphetherphe 650
 5594 AAGCATTCCTGCAGCAAAATTTGGTGAAGCTGGCATATTTGTTTCTCTTT 5643
 651 TyrHisVal.....SerCysGlnGlnLysGln..... 659
 5644 GCCAGGTTGTTGAAGCTCAGTACGTGTGAGAGAACTCAAGCATTTAGCACT 5693
 659 659
 5694 GCGAGTTACTCTCATTAATTAATACACAACAAGAGATCATGAT 5743
 659 659
 5744 TGAAGAAATTTGAATGAGCTTTCTATGATTCATCAGGTGTGACAAACA 5793
 660 AsnThrProLeuGluThrProVal 667
 5794 AATGCAATTCGTGGGTTTACATTTTGAAGACCTCTTGAAGATGCTG 5843
 667 LcLysThrThrPlePrometLeuGlnAsnGlyArgLeuLysThrGlyG 684
 5844 TGTGAGAGATATTATTATATGCAATGAGAAAT.....GAGAG 5878
 684 LnPheCysLeuProValSerLeuGluLysProProGlnAlaLysSerAl 700
 5879 ACTTTAAGTTGATGATGAGCTTAATGCTATTAATCCAAAGATGTTAAGTGG 5928
 701 LeuSerProGluValProLeuProGluLysThrValAspAsnHisLys 717
 5929 TTAGAG...GAACATATGCTGCTGACGGAAGATGAGAGTAAGCAGACAG 5975
 717 scLysValPheAsnValGluValAlaValSerSerLeuHisThrGlnA 734
 5976 AGGTGTGGGAACTTGTGCTAGCAGCTGTAGAACTCTCATCAGAGCAG 6025
 734 sPProTyrLeuAspLysPhePheAlaLeuValAsnAlaLeuAspLunHis 750
 6026 ATCACCACACGACAGATGTTT..... 6045
 751 LeuPheProValArgLLeuLysPheArgLLeuMetGluAsnAsnLeuG 767
 6046AATATTA 6053
 767 uAsnGluLeuLysSerSerLeu..... 774
 6054 GCAGTTATTTGAAAGCTCAAGTGTGCATCAGTTCTACTGACTGTGTCAG 6103
 775SerAlaLeuAsnSerSerGlnLeuGluProValValArg..... 787
 6104 TTTTGCAGAAATACAAAGAGGGGACATCAGACCCATGCCCCGAGAGGTT 6153
 788PheLeuHisLeuLeuLeuAspLysLeu..... 796
 6154 TGTAGATCATTTGTGAATAATTATACAGAGAGTCTTGGATCTCCSCAGA 6203
 797IleLeuLeuValIleArgPro 804
 6204 TTTGGAATTAATGCAAAATTAATTCATATTCCTTTAGCAGTTCAACCTC 6253
 804 roVal..... 805
 6254 CTACATAACTTACGTTTGTGCACAAATCCACAGAACTTCTACTTTTCTTGG 6303
 806 ...IleAlaGlyGlnIleValAsnLeuGlyGlnAlaSerPheGluAlaMet 821
 6304 CACATAGATGGGAGAGATCTTTCAG.....GAGAAAGT 6335
 821 tAlaSerLeuIleAsnArgLeuHisLysAsnLeuGluGlyAsnHisAspG 838
 6336 GCGGTCATTCATGTACTCTG.....AGGCATTTCA 6364

838 LHisGlyArgAsnSerLeuLeuAlaSerTyrTLeuHisTyrValPheArg 854
 6365 GCACTGGAGGAGGTCCTTATAGACCTT.....GCATTTATGTAATA 6408
 855 LeuProAsnThrTyrProAsnSerSerProGlyLysGlyLeuG 871
 6409 AGCCCATCTGCTTTT.....ACTGCTTCACCATATGAAGAGAGAAATTC 6452
 871 yGlySerValHisTyrAlaThrMetAlaArgSerAlaValArgProAla 888
 6453 CTCTAATATTATTCCACAAACAGATGCCGCCCATATGCTGCTGTAGAA 6502
 888 erLeuAsnLeuAsnArgSerArgSerLeuSerSerSerAsnProAspIle 904
 6503 GCCTACCAAGCATTCCTACTCTTCACTACTAAGCAATCAACAAACAGTG 6552
 905 SerGlyThrProThrSerProAspAspGluValArgSerIleLeuLys 921
 6553 ACTGGAAGTTGGGTGTGTACTATCGACAGGTTCACAAATATT..... 6594
 921 rLysGlyLeuAspArgSerAsnSerTyrValAsnThrGlyLysProLysA 938
 6595GCAGATACTTATATGTTGCCACCAATCAAGAAAGAAC 6628
 938 lAlaLysProLysSerAsnProSerProSerAlaGluSerThrGlnAla 954
 6629 AAAATTCCTTTGGGAGTCCGACACACTGMAAAAGCAAGAGAGCGCA 6678
 955 MetAspArgSerCysAsnArg..... 961
 6679 TTCAATCAGTACGTGTGAGCTGCACAAACCTGTTTGGAAATGCAACTGT 6728
 962 .MetSerSerHisThrGluThrSerSerPheLeuGlnThrLeuThrGlyA 978
 6729 CCTCTACGCCCAAGTCTGTGTCAGTATGCCCAAGAGGAGCTCTGGGA. 6777
 978 rGluProThrLysLysLeuPheHisGluGluLeuAlaLeuGlnTrpVal 994
 6778 .TTTCCAGTGGTCMAAGCAGATCATTAACAGTTGGAGCAGAA..... 6819
 995 ValCysSerGlySerValArgGluSerAlaLeuGlnGlnAlaTrpPheP 1011
 6819 6819
 1011 ePheGluLeuMetValLysSerMetValHisHisLeuTyrPheAsnAspL 1028
 6820CCACAGCTCAG 6829
 1028 rLysLeuGluAlaProLysArgLysArgPheProGluArg.....PheMet 1042
 6830 AAGATGACATCTCTGGGAGTGAATCTCTGCCACGCGGACCTGATTACTTA 6879
 1043 AspAspIleAlaAlaLeuVal.....SerThrIleAlaSer..... 1054
 6880 AAGGAGATGGCTCTCTCCAGGACGACCAACAGCAGCATTTGCAAGCTTGG 6929
 1055A 1055
 6930 GCTAGCTTTTCTTCCACAGAACGAGATCTGACGCTTGGCCGTTGCCAA 6979
 1055 spLysValSerArgPheGlnLysAspThrGluMetValGluArgLeuAsn 1071
 6980 GTCTGTGTGATAGAAACCTGATGATTGGAA..... 7011
 1072 ThrSerLeuAlaPhePheLeuAsnAspLeuLeuSerValMetAspArgL 1088
 7012 ...AACTTTCCTATTCCTCTGTGTATGAGCCAAATTAACCGAAGTCTG 7058
 1088 yPheValPheSerLeuLysSerCysTyrLysGluValSerSerLysL 1105
 7059 AAGTGTCAACAGTGAAGTGAAGATGTGTGTGTAACCTATATGCTGTGAT 7108


```
8501 ATGAGAGAAAAAGTGATCCAGCATATTAGAGAAATGATATAAGTAGAT 8550
1648 ValIysMetLysGIuHISGIuAsPProGIuMetLeuIleAsPLeuMe 1664
      ::::: ::::: ||| ::::: ||| ::::: |||
8551 TTGAGTCCAGCAGCATGTGGCAG....GAACTTATTTCAGCAGCTGCAC 8594
1664 tTyraGIleAlaLysGIuTYrGIuThrSerProGIuArgLeuThrTP. 1680
      ::::: ||| ::::: |||
8595 ACATGATAGAGCAGTATGTRATGACCCCATCTACTATCCAACTCATGGC 8644
1681 .LeuGIuAsnMetAlaGIuLysHISerGIuArgSerAsnHISAlaGIu 1696
      ||| ::::: ||| ::::: |||
8645 AGTTGGATCCAAAGAGGCCAAATCGAGAGAG.....AGACGT 8685
1697 AlaAlaGIuCysLeuVal..... 1702
      ::::: ||| ::::: |||
8686 TTACAGAGATGTATTAACTATTCCAAATAGATCTCCTTAGGATAG 8735
1703 ...HisSerAlaAlaLeuValAlaGIuTYrLeuSerMetLeuGIuAsPa 1718
      ||| ::::: ||| ::::: |||
8736 ACAGAAATCAGAGATGTGTCAAAACCAACCACTCTTACTGTTTGAG 8785
1718 rGIuTYrLeuProValGIuCysValThrPheGIuAsnIleSerSerAsn 1734
      ||| ::::: ||| ::::: |||
8786 ACAA.....ACTCATCTCTTCTCTTCTTCTACT 8814
1735 ValLeuGIuGIuSerAlaValSerAspAspVal..... 1745
      ||| ::::: ||| ::::: |||
8815 GTC...AAAGACAAAGCTGCAAGTGAACTATTAGAGTGAATCGAGAGATG 8861
1746 .....ValSerProAspGIuGIu.....GlyI 1753
      ||| ::::: ||| ::::: |||
8862 CATCAGTGTTCACCATCTAGACAGACAGCTGTGGAATGTACTAGTA 8911
1753 IeCysSerGIuLysTYrPheThrGIu..... 1761
      ||| ||| ||| ::::: |||
8912 AATGT...GGAATGTATTGTGTGAAGATATGCTTGATACAGTTGAA 8958
1762 ...SerGIuLeuValGIuLeuLeuGIuAlaAlaIleSerPheSerMe 1777
      ||| ::::: ||| ||| ||| ::::: |||
8959 AGTTCGAGCCTTCAGGAGAGAGTGGAAACCAAGCATCATTTCTCGACA.. 9006
1777 talagIyMetTYrGIuAlaValAsnGIuValTYrLys 1789
      ||| ||| ::::: ||| ::::: |||
9007 .....TATGAAGAATTAAGAAGTTCACAG 9033
```